UNIVERSITY OF HELSINKI

Department of Food and Environmental Sciences

DEVELOPMENT AND APPLICATION OF HIGH THROUGHPUT SYSTEM FOR PROCESSING FECAL SAMPLES FOR MICROBIOTA ANALYSIS: PILOT STUDY OF PREVALENCE OF *Lactobacillus rhamnosus* GG IN INFANTS AND MOTHER-INFANT PAIRS

Prabin Koirala

Helsinki

May 2018

Tiedekunta/Osasto Fakultet/Sektion – Faculty
Agriculture and Forestry

Laitos/Institution– Department
Food and Environmental Biosciences

Tekijä/Författare – Author
Prabin Koirala

Työn nimi / Arbetets titel – Title
Development and application of high throughput system for processing fecal samples for microbiota analysis: pilot study of prevalence of *Lactobacillus rhamnosus* GG in infants and mother-infant pairs

Oppiaine /Läroämne – Subject
MBIOT (Microbiology and Environment)

Työn laji/Arbetets art – Level	Aika/Datum – Month and year	Sivumäärä/ Sidoantal – Number of pages
M. Sc. Thesis	May 2018	88

Tiivistelmä/Referat – Abstract

The current multidisciplinary interests on human intestinal microbiomes have stimulated large scale research initiatives, involving collection and processing of up to thousands of fecal samples within a single study. Hence, there is a need for high throughput protocols that are cost-efficient and validated for their performance to ensure that the relative abundance of different bacteria, the main outcome of microbiota studies, is not biased due to technical artefacts originating from sample processing. Infant's microbiota colonization is one of the central research areas in human microbiome research because of the long-lasting and profound health implications of the pioneering microbes.

This experimental study aimed to develop and validate a high throughput fecal sample collection and processing system for microbial DNA extraction operating in 96 well format. This newly developed method was used to extract DNA from 647 fecal samples collected from mother-infant pairs within a clinical study that will study the effect of antenatal antibiotic prophylaxis on infant's gut microbiota development. A subset of 28 mother-infant pair samples (14 from each antibiotic and non-antibiotic groups) were selected to study the prevalence of a probiotic bacterium, *Lactobacillus rhamnosus* GG (LGG), among infants and their mothers longitudinally from birth to 3 months by using species-specific PCR amplification method targeting sortase C gene. In addition, the prevalence of *L. rhamnosus* GG in 3-month-old infants was compared between the above samples and those (n=30) collected in another clinical trial conducted ~10 years earlier.

From extensive testing and validation, an efficient high throughput system for fecal sample collection and processing for extraction of microbial DNA in 96 well format was established. Tests were performed to validate the performance of a) fecal sample collection system b) commercial, readymade bead beating tubes for bacterial cell lysis and c) self-made wash buffers as part of the automatic DNA purification system. Performance was evaluated based on the quality and quantity of the resultant DNA. We show that this new fecal processing system can yield high quality microbial DNA from 96 fecal samples within ~6 hrs. Based on the ratios of dominant gram-positive and gram-negative bacteria evaluated using PCRs and next generation sequencing, the new DNA extraction method resulted into similar microbiota composition than the previously validated manual DNA extraction method. However, the DNA yield per sample was markedly lower due to the lower input volume of the sample. Based on the sortase C gene PCR tests, the prevalence of LGG was similar (~60%) among 3-month-old children in both clinical studies conducted ~10 years apart, although false negatives among the recent samples due to the low amount of DNA cannot be excluded. Following the temporal pattern of colonization, we observed no evidence for the transfer of LGG at the time of birth from the mother to her child, instead the infants became positive for LGG typically between 1-3 weeks after birth. The carriage of LGG seemed to be dependent on their diet. During this project, we found out that the PCR method employed for detection of LGG was not fully specific for this strain, and hence a more specific qPCR assay was developed.

Avainsanat – Nyckelord – Keywords
Gut microbiota, colonization, fecal samples, DNA extraction, high throughput, automatic purification, PCR, *Lactobacillus rhamnosus* GG, LGG, sortase C, bead beating, probiotic

Säilytyspaikka – Förvaringställe – Where deposited
The Digital Repository of University of Helsinki, Helda, Viiki Science Library

Muita tietoja – Övriga uppgifter – Additional information

PREFACE

The basis for this research originally emerged from the practical need for high throughput method to process thousands of fecal samples for gut microbiota analysis in the context of Health and Early Microbiota (HELMi) project and my curiosity in understanding the excellent probiotic property of *Lactobacillus rhamnosus* GG. This thesis project was carried out in the research group of Professor Willem M. de Vos and Docent Anne Salonen at Department of Bacteriology and Immunology, University of Helsinki in 2016/17 under the supervision of Anne Salonen. I would like to express my gratitude to Anne for her friendly behavior and support throughout the project. I have always learned new things from her and thank you for being at every obstacle that we have faced and solved together. Your generosity has encouraged me to further pursue my career towards research.

My sincere thanks to Professor Willem de Vos for providing me an opportunity to participate in his research and would also like to appreciate his insights and humbleness that has helped me to explore new area of gut microbiota and its influence on human health.

Special thanks to Anne Nikkonen for her invaluable effort in collecting, managing and providing me samples from participants of JORVI study for this project.

My special thanks to Pia Rasinkangas and Laura Huuskonen who have helped me to learn new techniques and have always provided insights for different projects. We were involved in different projects conducted in labs at Viikki and Haartman. Everyone from our group and Reetta Satokari research group, thank you for your support in executing different experiments and making lab a wonderful place. I would also like to thank everyone in the immunobiology research program for their support in completion of my project.

Finally, I am thankful to my parents and all my friends who have always supported me in every step of my life and studies.

Prabin Koirala

Helsinki

May 2018

Table of contents

ABBREVIATION

Cq	Threshold Cycle
DNA	Deoxyribonucleic acid
EDTA	EthyleneDiamineTetraaceticAcid
EtOH	Ethanol
HMP	Human Microbiome Project
HWI	Self-made wash buffer I
HWI_E	Self-made wash buffer I with EDTA
HWII	Self-made wash buffer II
IBD	Inflammatory Bowel Disease
IBS	Irritable Bowel Syndrome
LGG	*Lactobacillus rhamnosus* GG
MetaHIT	Metagenomics of the Human Intestinal Tract
Na-Ac	Sodium Acetate
NGS	Next Generation Sequencing
PCR	Polymerase Chain Reaction
qPCR	Quantitative PCR
RBB	Repeated Bead Beating
RNA	Ribonucleic acid
rRNA	Ribosomal RNA
RT	Readymade bead beating tubes
srtC	Sortase C gene
ST	Self-made bead beating tubes
Tris-HCl	Tris Hydrochloride

1. Introduction

As soon as infant in the womb is born, natural colonization of the microbiota takes place through contacts with mother's vagina, excreta and skin (Funkhouser and Bordenstein 2013; Mahony et al. 2018). These environmental exposures are the foundation for establishing a symbiotic microbial community, which matures during childhood and from that on remains relatively stable (Trasande et al. 2013). Microbiome and microbiota are the two terms used to address the host-associated microbial communities. The term microbiome was first coined by Joshua Lederberg in 2001 (Lederberg and McCray 2001), since then it has become popular among researchers. Although strictly speaking the term microbiota refers to microbial species living on or in human body and microbiome additionally to the complete set of their genes, presently these terms are used interchangeably (Turnbaugh et al. 2007).

Microbial cells outnumber the total cells in human body, most of them residing in the gastrointestinal tract (Ursell et al. 2012; Sender, Fuchs, and Milo 2016) from where they contribute to key physiological processes, such as digestion, immunological maturation and modulation and even behaviour (Sekirov et al. 2010; Qin et al. 2010; Ursell et al. 2012; Cox et al. 2014; Tanaka and Nakayama 2017). Hence, scientists are now focused in defining the symbiotic relationship of human and the microbiota (NIH HMP Working Group et al. 2009). Accumulating evidence indicates that a child's healthy future is dependable on early colonization and its effects on immunological and other physiological programming (Fisher et al. 2017; Jenmalm 2017), and the factors driving early microbiota development and the resultant influences on child's health are being studied extensively (Mueller et al. 2015a). Numerous factors are responsible for shaping the human microbiota starting from the mode of birth (Rodríguez et al. 2015). The two to five first years of life is the most critical period to establish a healthy microbiota towards an adult-like composition (Koenig et al. 2011; Yatsunenko et al. 2012; Borre et al. 2014); disruption in natural microbial colonization and succession can lead to increased risk of chronic diseases later in life (Rodríguez et al. 2015; Funkhouser and Bordenstein 2013; Rautava 2017).

The intestinal microbiota acts as a physical barrier against pathogenic microbes and perform some of the important metabolic activities, such as digestion and nutrient metabolism (mostly enzymatic) and drug metabolism. Our immune system can develop tolerance and function appropriately from the beginning and throughout the lifetime with the help of microbes residing

in our gut (Jandhyala et al. 2015). An enormous firewall system is responsible in maintaining homeostatic relationship between the microbes and the host, and breakdown in this homeostasis can result in several inflammatory conditions and diseases (Macpherson et al. 2009; Belkaid and Hand 2014; Isolauri, Sherman, and Walker 2017; Fung, Olson, and Hsiao 2017; Rowland et al. 2017; Tanca et al. 2017).

Feces is the most feasible and popular material to study the gut microbiota, representing the microbial population residing in the lower gastrointestinal tract (Matsuki et al. 2002). Culture-dependent techniques are not suitable for large-scale microbiota studies due to their selection bias and labour-intensity. Molecular, culture-independent techniques are popular nowadays as they can successfully address these problems and can be adjusted to fulfil the required criteria (Hiergeist et al. 2015). Next Generation Sequencing (NGS) is currently the most popular approach for high throughput microbiota analysis, typically based on the amplification and sequencing of the 16S rRNA gene (Eckburg et al. 2005; Schuster 2008; Wang et al. 2009; Jandhyala et al. 2015; Walker et al. 2015). NGS has successfully replaced culture-dependent (Suau et al. 1999; Rajilić-Stojanović and de Vos 2014) old methods and when compared to Sanger sequencing, this technique provides massive parallel sequencing in a cost-effective way (Sanger, Nicklen, and Coulson 1977; Behjati and Tarpey 2013). Following the first large scale microbiome projects, the US Human Microbiome Project (HMP) and European Metagenomics of the Human Intestinal Tract (MetaHIT) (Qin et al. 2010; Human Microbiome Project Consortium 2012), the number and volume of human gut microbiota studies have increased exponentially with recent studies analysing up to thousands of samples (Beaumont et al. 2016; Falony et al. 2016). Hence there is an increasing need for efficient HTP sample processing methods for molecular microbiota analyses based on fecal samples. An efficient fecal sample processing system requires an appropriate sample handling protocol and a robust and efficient DNA extraction method. Mechanical lysis of the cells has been shown to be essential for efficient lysis of both gram positive and negative bacteria to gain a good representation of the full diversity and species composition of the gut microbiota (Salonen et al. 2010; Smith et al. 2011; Claassen et al. 2013; Walker et al. 2015).

In this MSc thesis a new high throughput system was tested and validated for extracting bacterial DNA from 96 fecal samples per run. The protocol was optimized for simple and complex bacterial communities from infants and adults, respectively. With this new system, based on the Repeated Bead Beating (RBB) and automatic purification methods, large number of samples can be processed in short period without compromising the efficiency of cell lysis.

The high throughput system was further optimized to improve both the quantity and quality of microbial DNA using self- made buffers for lysis, washing and elution instead of those available in the extraction kit, to optimize the DNA preps for PCRs, sequencing and other downstream experiments. Optimized high throughput system was then used to extract microbial DNA from 634 fecal samples of mother-infant dyads, summing up to 342 adults and 292 infant samples, represent different time points since childbirth. A subset of the samples was analysed with PCR and qPCR to quantify the probiotic *Lactobacillus rhamnosus* GG (LGG) to determine its carriage rate as well as the age at which the infants became positive for LGG. In addition, the prevalence of LGG was studied in 3-month-old children from two different birth cohorts sampled ~10 years apart.

2. Literature Review

2.1. Use of feces as study material

Fecal samples provide a non-invasive material for the analysis of the gastrointestinal, predominantly colonic, microbes (Biasucci et al. 2008). The gastrointestinal microbiota consists of up to hundred trillion bacteria, which significantly contribute to our health and well-being by food digestion and nutrition intake, neutralizing pathogens, and shaping our immune systems (Tappenden and Deutsch 2007; Belkaid and Hand 2014). Development of the human microbiota throughout the lifetime, their functions and study approaches are being discussed under this section.

Current intestinal microbiota studies are largely dependent on the fecal samples that can be collected, stored and processed using different methods. Feces can unravel wide range of information from individuals and colonic luminal microbial composition excluding the mucosal layer and small intestine which require more invasive approach (Zoetendal et al. 2002; Eckburg et al. 2005; Aguirre and Venema 2015). Comparatively, convenience in collecting fecal samples without intricacies makes it very popular and useful method for studying the human gut microbiota. Briefly, feces is composed of 75% water and 25% solid material (30% deceased bacteria, 40% proteins, fat and inorganic substances and 30% undigested foods and other components) (Barbosa 2013).

Collection, storage and processing of fecal samples especially for molecular studies have always challenged researchers, as several factors comes in to play that can affect the outcome of study (see next section). For an optimal high-throughput system, sample handling time should be minimized whilst maintaining the quality of samples. Weighing the required amount of sample is the most tedious work which takes half of the total processing time and eliminating this step to process fecal samples for DNA extraction was our priority. Child fecal samples are normally semi solid or watery and messy so it's trickier to collect baby fecal samples compared to adults. Thus, a standard fecal sample processing system is need for reproducible results. A new high throughput fecal sample processing system, starting from the collection of feces to quantification of extracted DNA was developed in this project to enable efficient processing of thousands of fecal samples for microbiota analysis.

Large scale study designs where thousands of fecal samples need to be collected set special requirements for sample collection, storage and transportation. All these steps have very important role in preserving the original ratio of microorganisms in the sample because the relative abundance of different bacteria is typically the main outcome of microbiota studies. Immediate freezing is considered as the gold standard in sample collection and thawing should be avoided before the DNA extraction (Tedjo et al. 2015; Costea et al. 2017). Shaw et al reported that infant's fecal samples represented original microbial community independent of their storage at -80°C and method of DNA extraction (Shaw et al. 2016). Lauber et al's study also revealed no compositional microbiota differences between 3 days and 2 weeks old samples stored at temperatures +20, +4, -20 and -80°C (Lauber et al. 2010). Finally, Tedjo et al's study also observed no significant differences among the total microbiota and number of species in fecal samples stored at different temperatures, whether from healthy subjects or unhealthy subjects (patients with Irritable Bowel Syndrome and Inflammatory Bowel Disease) and independent of fecal consistency (Tedjo et al. 2015). Although these studies indicate that the fecal microbiota composition is relatively stable during storage and the individual-specific profiles are not overridden by technical factors, aliquoting and storing at -80°C while eliminating thawing before first DNA extraction is considered as the best option to ensure that technical variation does not mask the more nuanced microbiota differences that may be important. These are also the recommendations of the FP7-funded International Human Microbiome Standards (IHMS) project (http://www.microbiome-standards.org/) that was established to develop standard operating procedures (SOPs) to optimize data quality and comparability for the microbiota studies on human fecal material.

2.2. Current methods used to study gut microbiota using fecal material

2.2.1. Microbial DNA extraction and purification

Modern human microbiome research is largely based on DNA extracted from fecal samples and hence the quality of DNA is crucial for the success of experiments and their outcome.

DNA extraction methods have evolved drastically as compared to the traditional phenol-chloroform method, which mainly focuses on removal of polymerase chain reaction (PCR) inhibiting components, their carryover to the next step (McOrist, Jackson, and Bird 2002), and importantly lacks throughput capability while also being unsafe. Traditional, time consuming additional clean up procedures have made scientists to think about using the solid matrix (silica) as a binding agent for the extracted DNA, which requires fewer washing steps to remove the undesirable components (Boom et al. 1990) and has been proven very effective in microbial DNA extraction from feces (Jones et al. 1993). Processing large number of samples using the above two methods have proven laborious, time consuming and costly, so the development and use of commercial kits has become popular. Several papers have compared different fecal DNA extraction methods and have shown that successful DNA extraction is dependent on cell lysis step which can rupture cell walls from both gram positive and negative bacteria as otherwise the sample is not representative for all bacteria present in the sample (Fujimoto, Nakagami, and Kojima 2004; Salonen et al. 2010; Santiago et al. 2014; Costea et al. 2017). Based on these studies. mechanical bead beating method has been shown to be efficient in rupturing the cell walls, especially of gram positive bacteria. The further steps for purification of the DNA mainly affect the yield and purity as well as the time needed to complete the protocol (Salonen et al. 2010). The RBB method that was used as a starting point in this MSc thesis is widely used and has been shown to perform well in extracting DNA from a wide range of bacteria found in fecal samples (Salonen et al. 2010; Costea et al. 2017, 1069).

As traditional methods can only handle limited number of samples at a time, researchers are now interested in high throughput extraction and purification systems to conduct large scale research. These new high throughput methods should also be efficient in producing high quality DNA, decrease the amount of labour and should be inexpensive. Traditional extraction and purification methods require multiple centrifugation steps which has been the major hurdle

for making the whole process automatic, to tackle this problem modern nucleic acid purification systems use paramagnetic beads. Use of paramagnetic beads (Figure 1) with modified surface capable of efficient DNA binding (and releasing by magnetic disconnection) have become popular in delivering high quality nucleic acids with minimal use of washing buffers (Lesley 2001; Akutsu et al. 2004). Use of magnetic beads for the separation enables extraction of nucleic acids directly from crude samples (feces, blood, culture media, homogenized tissues, etc.) whilst eliminating the carry-over of biological debris and the need for extra laborious steps required to pre-process the samples. All these features make this method well suited for high-throughput DNA extraction and sample processing (Šafařík, Ptáčková, and Šafaříková 2001; Franzreb et al. 2006).

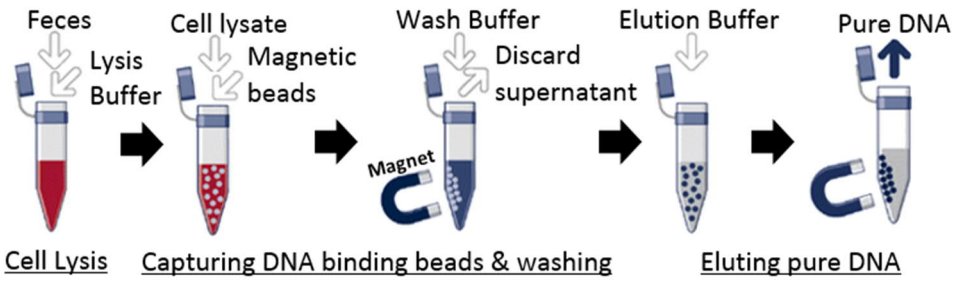


Figure 1 Illustration of basic principle of nucleic acid purification using paramagnetic beads (figure adapted from Creative Diagnostics® and principle from Thermo Scientific™ KingFisher™)

2.2.2. Polymerase Chain Reaction (PCR)

PCR is a valuable tool for studying gene expression, identifying and quantifying pathogens, and conducting clinical diagnostics including forensics. It has become a starting point of many advanced molecular techniques and has been used extensively since its development (Weissensteiner and others 2003). In a normal end-point PCR, oligonucleotide primers specific to the targeted DNA is used to amplify the given DNA segment and the results are visualized in agarose gel to evaluate the intensity and size of the amplicon(s). With many advantages, end-point PCR lacks real time visualization and quantification of amplifications and there is

always chances of getting false amplifications. However, end-point PCR is extensively used in detection of various bacteria using ribosomal RNA (rRNA) or other bacterial genes as targets from different sample types including feces (Picard et al. 1992; Matsuki et al. 2002; Wang et al. 2009). For example, specific primers were used to amplify and detect *Bifidobacteria* and *Lactobacillus* species from infants' and adults' fecal samples (Kok et al. 1996; Walter et al. 2001), and sortase C gene-targeted end-point PCR was used to differentiate a specific probiotic strain *Lactobacillus rhamnosus* GG (LGG) from other *Lactobacillus* strains and probiotic bacteria (Vélez, De Keersmaecker, and Vanderleyden 2007; Kankainen et al. 2009; Call and Klaenhammer 2013; Douillard et al. 2014; Rasinkangas et al. 2014).

2.2.3. Real Time PCR or quantitative PCR (qPCR)

Unlike the normal PCR, qPCR enables real time monitoring of nucleic acid amplification and quantification of the products. Firstly introduced by Holland et al., it has been continuously developing since and with the introduction of new fluorescent dyes this method has become more sensitive and reliable in targeting and quantifying specific nucleic acids (Holland et al. 1991; Lee, Connell, and Bloch 1993; Heid et al. 1996; Wall and Edwards 2002). For human microbiome research, qPCR has been used extensively for detection and quantification of specific bacteria in fecal samples (Olsen et al. 1986; Vaughan et al. 1999; Blaut et al. 2002; Matsuki et al. 2002; Wang et al. 2009; Sekirov et al. 2010; Walker et al. 2015) while more recently the use of this targeted method has decreased due to the advanced sequencing techniques providing relative abundance of all detected bacteria. However, qPCR is still the method of choice when absolute quantification of certain bacteria is desired. In this thesis project, a new qPCR assay was developed to replace the end-point PCR on *L. rhamnosus* GG sortase C gene that lacked the specific identification, quantification and throughput properties needed to correctly identify and quantify this specific probiotic species in fecal samples.

2.2.4. Next Generation Sequencing (NGS)

NGS is a revolutionary technology that has modernised genomic research by enabling DNA sequencing in high throughput mode. NGS possesses the capability of massive parallel sequencing with unprecedented speed and accuracy, and have become the workhorse of human microbiome studies (Sanger, Nicklen, and Coulson 1977; Schuster 2008; Pareek, Smoczynski, and Tretyn 2011; Van Dijk et al. 2014; Reuter, Spacek, and Snyder 2015; Daliri et al. 2017; Lyons et al. 2017). Several NGS platforms are available and they use different sequencing technologies (Behjati and Tarpey 2013). In microbiome studies, Illumina's MiSeq and HiSeq platforms are widely used, most typically for sequencing of the 16S rRNA gene amplicons that provide information on bacterial composition and diversity after bioinformatic analysis.

2.3. Gut microbiota in infants and adults

Following birth, infant's gut is colonized by diverse microorganisms. The first colonizers are facultative anaerobes, such as *Enterobacteriaceae* and Bacilli, and within the following days and weeks, anaerobic *Bifidobacterium, Clostridium* and *Bacteroides* start to dominate (Matamoros et al. 2013). *Bifidobacterium* spp. are specialized in fermenting human milk oligosachharides and thrive for the first few months especially in breast-fed babies. With the introduction of solid foods their numbers start to decrease and *Bacteroides, Ruminococcus* and *Clostridium,* specialized in metabolizing complex dietary polysaccharides take over (Koenig et al. 2011; Bäckhed et al. 2015) and by the age of 2 to 3 the microbiota largely resembles that of an adult. During child's gut microbiota's development period, extrinsic factors have more effect on the development of gut microbiota than intrinsic factors (Fallani et al. 2010; Koenig et al. 2011; Munyaka, Khafipour, and Ghia 2014). A recent study has also reported the exceeding impacts of environmental factors over genetics in shaping the adult gut microbiota (Rothschild et al. 2018).

In adulthood, the gut microbiota is remarkably stable and the subject-specific composition persists at least for a decade, presumably longer (Rajilić-Stojanović et al. 2013). There are substantial differences in microbial composition between individuals at species and genus levels but practically always the dominant ones belong to *Clostridium* clusters XIVa and IV

among Firmicutes or *Bacteroides spp.*, reaching up to 60%-80% of total bacteria depending on the individual (Hayashi, Sakamoto, and Benno 2002; Zhernakova et al. 2016; Lyons et al. 2017). *Proteobacteria* and *Actinobacteria* typically make up to 5%-10% of total bacteria (Wang et al. 2003; Eckburg et al. 2005; Agans et al. 2011; Lay et al. 2005). Most adults (80%) are positive for *Lactobacillus* but like the other facultative anaerobes, e.g. *Eschericia coli, Enterococci* and staphylococci, they normally represent less than 1 percent of adult's fecal microbiota (Adlerberth and Wold 2009).

2.4. Factors affecting the development of infant's gut microbiota

The microbial colonization starts at or after birth and children's microbiota undergoes succession with changes with the environmental and endogenous conditions (Stark and Lee 1982; Benno 1985; Mueller et al. 2015a; Rodríguez et al. 2015; Tanaka and Nakayama 2017). There are several factors (Figure 2) that can affect the normal colonizing process in infants and overall modify the microbiota (Guarner and Malagelada 2003; Funkhouser and Bordenstein 2013).

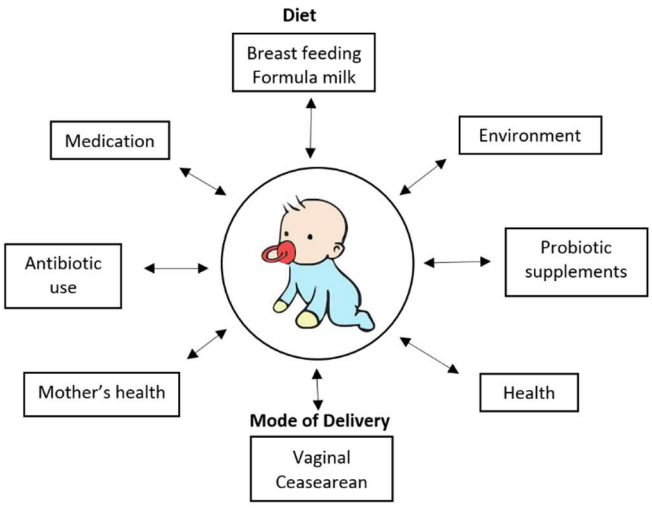


Figure 2 Factors that can affect infant's gut microbiota development and influence the microbiota overall.

Several premature and term birth studies have suggested the presence of microorganisms already in the placenta and meconium, which points out the scenario that infants may get first introduced to microbes already inside the womb (Aagaard et al. 2014; Ardissone et al. 2014), through placental barrier (Aagaard et al. 2014), through umbilical cord blood (Jiménez et al. 2005), through amniotic fluid (Bearfield et al. 2002) and fetal membrane (Rautava et al. 2012) without development of any kind of inflammation or infection. Staphylococci, streptococci and lactobacilli are the dominant bacteria in meconium (Moles et al. 2013; Hu et al. 2013).

The first week after the birth child's microbiota is very similar to mother's skin and vaginal flora (Stark and Lee 1982; Koenig et al. 2011). The mode of delivery (vaginal vs. caesarean) is regarded as one of the main factors shaping the early microbiota. The caesarean delivered babies are lacking normal colonizing pattern, especially *Bifidobacteria* and *Bacteroides* spp. for months when compared to vaginally delivered children (Bäckhed et al. 2015; Azad et al. 2016). Naturally delivered infants are first exposed to the bacterial species most prominent in the vagina and feces of the mother rather than those born by caesarean method who lack these maternal exposures (Biasucci et al. 2008).

Diet is one of the most influencing factors in shaping the human microflora throughout the lifetime. Earliest dietary effects on the gut microbiota are due to the introduction of mother's milk versus formula feeding, while later the intake of cereals, vegetables and other foods with higher fiber content associate with the microbiota composition and their diversity (Biasucci et al. 2008; De Filippo et al. 2010; Walker et al. 2011). Human breast milk is the first and most important food of an infant that among other benefits helps in the development and maturation of the gut microbiota and is recommended as an exclusive nutrient source for the first six months (Kramer and Kakuma 2004; Cattaneo et al. 2005; Section on Breastfeeding 2012). Introduction of breast milk provides the infant with different types of growth factors, antimicrobial peptides and growth hormones along with microorganisms and their preferred growth substrates that help in maturation of the intestinal microbiota and immune system including the gut barrier and defence systems (Oftedal 2002; Petherick 2010; Walker 2010; Bezirtzoglou, Tsiotsias, and Welling 2011; Ballard and Morrow 2013). Alteration in natural colonization of the gut microbiota during infancy has been associated with an increased risk for inflammation, autoimmune diseases, allergies, diabetes, Inflammatory Bowel Disease (IBDs), obesity and atopic dermatitis (Mackie, Sghir, and Gaskins 1999; Akobeng et al. 2006; Ogra and Welliver RC 2008; Guaraldi and Salvatori 2012).

Inclusion of probiotics as supplement and in staple foods have become popular as they can trigger notable health benefits especially in diarrhoeal diseases and potentially in preventing allergy (Metchnikoff and MITCHELL 1907; Hotel and Cordoba 2001; Pflughoeft and Versalovic 2012). Extensive use of probiotics can alter the amount of microbial commensals in gut, suppress pathogens, stimulate the proliferation of epithelial cells and fortify intestinal barriers in both infants and adults (Othman, Neilson, and Alfirevic 2007; Thomas and Versalovic 2010; Bron, Van Baarlen, and Kleerebezem 2012; Hemarajata and Versalovic 2013; Govender et al. 2014). *Lactobacillus rhamnosus* GG, originally isolated from human gastrointestinal tract, is one of the most studied probiotic bacteria which has many health benefits (Alander et al. 1999; Gorbach 2000; Näse et al. 2001; Doron, Snydman, and Gorbach 2005; Parvez et al. 2006). Various studies have revealed *L. rhamnosus* GG's effectiveness against respiratory tract infection and different types of diarrheas in children (Guandalini et al. 2000; Hatakka et al. 2001). Meticulous study of *L. rhamnosus* GG's proteinaceous host-interacting pili and pilin subunit SpaC have identified it to be the key component for interacting with host tissues and conferring the prolonged persistence of *L. rhamnosus* GG in human gastrointestinal tract which lacks in other strains of *Lactobacilli* (Kankainen et al. 2009; von Ossowski et al. 2010; Reunanen et al. 2012; Douillard et al. 2014; Rasinkangas et al. 2014).

Use of antibiotics is also one key factor that can influence the development of the microbiota. Use of antibiotics can have profound effects upon adult's microbiota but especially during the vulnerable microbial colonization and maturation process in children (Van der Waaij 1983; Ledger and Blaser 2013; de Tejada 2014; Stokholm et al. 2014; Korpela et al. 2016; Langdon, Crook, and Dantas 2016). This disturbance in colonization process has been directly associated with infant's risk for obesity, diabetes, IBDs, atopic dermatitis and asthma during their childhood and later in life (Kozyrskyj, Ernst, and Becker 2007; Algert et al. 2009; Decker et al. 2010; Hviid, Svanstrom, and Frisch 2011; Huh et al. 2012; Trasande et al. 2013; Mueller et al. 2015b).

3. Aims and Objectives

Aim of this experimental study was to develop and validate a high throughput method for processing and extracting bacterial DNA for microbiota analysis from human fecal samples. The study consisted of two parts with the following objectives:

To develop, optimize and validate high throughput fecal sample processing system for gut microbiota analysis

- To develop a fecal sample collection method and storage system optimal for high throughput studies
- To optimise and validate an efficient, semi-automated and high throughput DNA extraction method for fecal samples, focusing on the optimization of cell lysis parameters and semi- automated DNA purification system to achieve cost-effective and high-quality DNA extraction

To apply the newly optimized and validated high throughput fecal sample processing system for extraction of DNA from fecal samples collected in a clinical trial and use the samples to study the prevalence of probiotic bacterium, *L. rhamnosus* GG in infant and maternal fecal samples using PCR methods

- To study the prevalence of probiotic *L. rhamnosus* GG in 3 months old infants from two cohorts sampled ~10 years apart
- To compare the transfer of *L. rhamnosus* GG within infant-mother dyads longitudinally

4. MATERIALS AND METHODS

4.1. Study design

This study was designed as master's thesis project to develop, validate and apply a semi-automated high throughput fecal sample processing system to extract microbial DNA. A subset of the samples was used to study the prevalence and carriage of *Lactobacillus rhamnosus* GG using species specific PCR amplification method. A detailed study design and workflow is displayed in Figure 3 below.

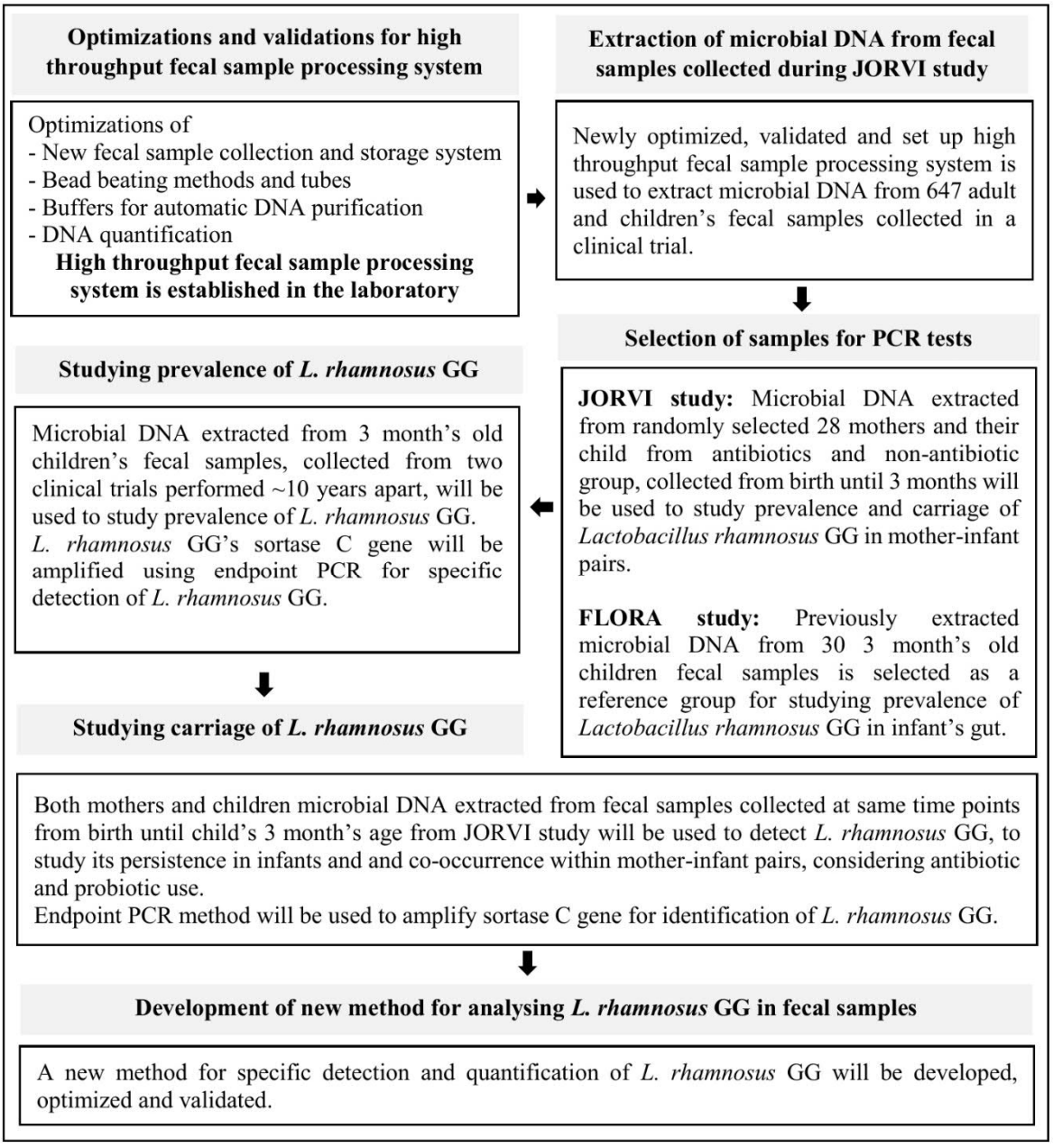


Figure 3 A brief explanation of this study designed as a Master's Thesis project

4.2. Trials and samples

Various in house and drop out fecal samples from infants and adults were used during the optimization and scale up of the fecal sample processing system. Selected samples had been previously processed and analyzed for the microbiota composition. Details of the samples are shown in Table 1.

Table 1 Fecal samples used for optimization and validation during the development of high throughput fecal sample processing and DNA extraction system

Sample ID	Group	Age (months)
Ab36	Child	5
Ab41		1
Ab63		2
Ab91		2
UH1	Adult	
UH2		
UH3		
INT1		

Fecal samples collected in a clinical trial involving mother and infants were processed using the developed method. Randomly picked subsets of these and previously extracted infant samples were used to study the prevalence and carriage of *L. rhamnosus* GG. These clinical trials are described below.

4.2.1. JORVI study

JORVI is an ongoing study led by Dr. Kaija-Leena Kolho from Helsinki University Hospital. It is a randomized controlled study where 60 mothers were recruited from the labour ward at Jorvi hospital, Espoo in 2015-2016. Half of the mothers (30) received prophylactic antibiotics during the labour and the rest (30) were not administered any antibiotics. Fecal samples from the mothers and infants were collected at seven different time points: At birth, 1 week (w), 2

weeks, 1 month (m), 3 months, 6 months and 12 months. The infant feces were collected from the diapers. In case of fully absorbed (liquid) feces, the sample was collected by tapping the diaper with cotton swab instead of a tube with scoop. Rectal swabs were taken from the mother at the day of childbirth if natural fecal sample could not be produced. Only samples collected by April 2016 were available for this study and hence were used for DNA extraction. By that time, 28 mother-infant pairs have donated the first 5 samples from birth until 3 months (Figure 3) which were used in this project to study the prevalence of *L. rhamnosus* GG among infants at 3 months and its co-occurrence between mothers-infant dyads at different time points using *L. rhamnosus* GG (Sortase C gene) specific end-point PCR. Figure 3 and Table 2 explain the JORVI study design and available data in brief.

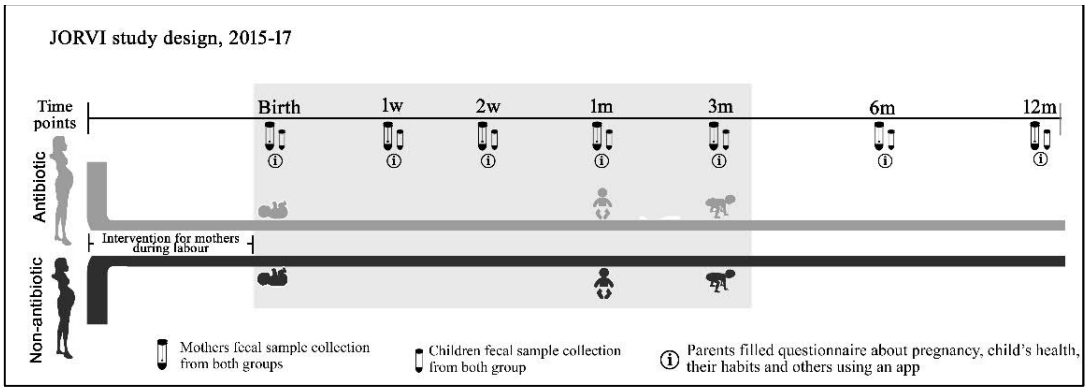


Figure 4 JORVI study design, first five samples (shaded) from birth until 3 months (3m) were used to study the prevalence of L. rhamnosus GG *in the samples.*

Table 2 Available information of 28 mother-infant pairs selected to study LGG's prevalence and its temporal dynamics. Fecal samples collected until 3 months and their relevant information were only used and analyzed for prevalence and study of transfer of LGG from pregnant mothers to their child.

Groups ➔ Headings ⬇	Mothers			Children	
	Antibiotics intervention until birth	Birth-3 months			
		Antibiotic	Non Antibiotic	Antibiotic	Non Antibiotic
Total	28	14	14	14	14
Average age (y.)	33.5				
Male				10	6
Female				4	8
Vaginal delivery	23				
Caesarean delivery	5				
Breast feeding				4	7
Formula milk				1	
Breast and formula milk				9	7
Under medication	6	4	3		
Additional antibiotics		2	4	1	2
Probiotic Supplement (PS)	6	7	5	13	12
PS containing LGG	5	5	2	10	6
Vitamin D supplement				11	11
Probiotic Foods (PF)	11	8	9		
PF containing LGG	5	4	1		
Drinking or using milk products	25	14	14		

4.2.2. FLORA study

FLORA is a probiotic intervention study conducted from 2003 to 2005 by Dr. Erkki Savilahti and Dr. Mikael Kuitunen from Helsinki University Hospital. Pregnant mothers whose infants had an increased risk for allergy were recruited at antenatal clinics through advertisements in Helsinki. Full description of the study and results regarding the effect of intervention on the incidence of allergic diseases at two years have been published in 2007(Kukkonen et al. 2007). In brief, pregnant mothers (N=1223) at 35 gestational weeks were randomized to either probiotic or placebo group. Mothers in the probiotic group received a capsule twice daily for 2 to 3 weeks before delivery containing a mixture of *Lactobacillus rhamnosus* GG, *Lactobacillus rhamnosus* LC705, *Bifidobacterium breve* Bb99 and *Propionibacterium freudenreichii* ssp.

shermanii JS (DSM 7076). Capsules contained microcrystalline cellulose in the placebo group. For 6 months after the birth, the infants in the probiotic group received an opened probiotic capsule with 20 drops of sugar syrup containing 0.8g of galacto-oligosaccharides daily. While the infants in the placebo group received sugar syrup without galacto-oligosaccharides. Fecal samples from infants were collected at 3 months and 6 months.

For this project, pre-extracted microbial DNA from 3-month samples from 30 children (born ~2003) were randomly selected from control group (placebo) to study the prevalence of *L. rhamnosus* GG in infant's gut against similarly analyzed fecal samples from 3 month's old children (born ~2015/16) from JORVI study. The following diagram explains FLORA study in detail (Figure 5).

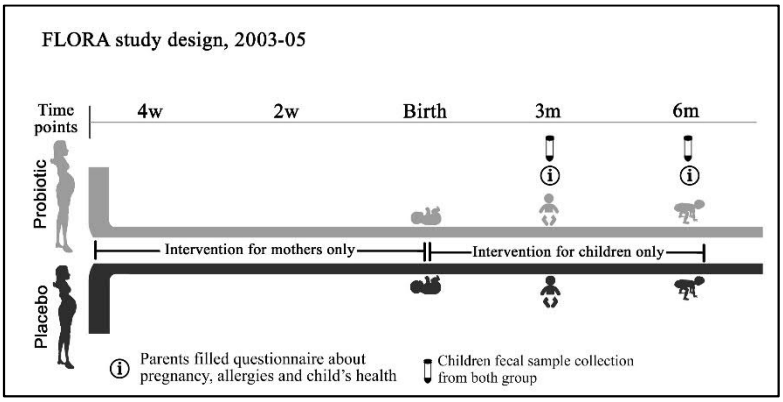


Figure 5 A detailed diagram showing FLORA study conducted between 2003 and 2005 in which half of the pregnant mothers and their child were administered probiotics 4 weeks before giving birth and after birth until 6 months, respectively. Gut microbiota is studied using fecal samples collected from the infants at age of 3 and 6 months. In this study, fecal samples from children in the placebo group at 3 months was used to study the prevalence of *L. rhamnosus GG* among 3 month's old infants.

4.3. Fecal DNA extraction and optimization of parameters

Setting up a new high throughput fecal sample processing system in 96 well format involves multiple steps and components. In this study, we tested, optimized and validated the following components and procedures as part of the workflow required to obtain PCR-compatible DNA from fecal sample (Figure 6).

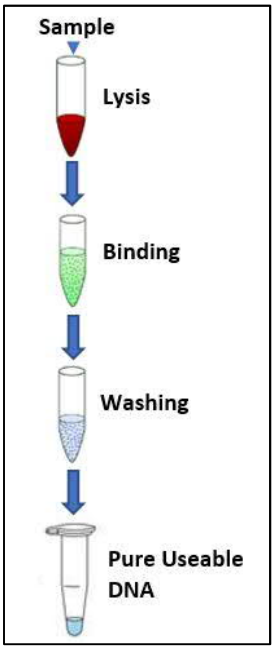


Figure 6 A simple workflow for obtaining purified DNA from fecal sample.

New fecal sample collection system for infants (sample), transit from self-made bead beating tubes to commercial ones and selection of beat beating conditions (lysis), transition from column-based (manual) DNA purification to automated process using KingFisher™ Flex (binding) and wash buffers for DNA purification (washing). In addition, minor experiments, not described in this thesis, were performed to upscale the bead beating from 24-format to 96-format, to accommodate 96 well plates to VorTemp™ 56 (Labnet International, USA) and to accurately quantify DNA in 96 well format. In the optimization phase bead beatings were performed in FastPrep®-24. A new protocol for extraction of microbial DNA from adult and infant fecal samples was established at the end.

4.3.1. Development of new fecal sample collection system for infants

Before optimizing DNA extraction and purification system, a new fecal sample collection technique was developed especially for infants according to the need in the forthcoming large HELMi-study focusing on early life microbiota. Four collection techniques were tested (Table 3) using real infant feces samples of different texture to obtain roughly ~250 mg sample, i.e. the required amount of fecal material for the manual repeated bead beating (RBB) method. Lab colleagues and two mothers were asked for their comments after testing different techniques. The newly developed infant's fecal sample collection procedure can be found in appendix A and properties of each technique is enlisted in Table 3.

Table 3. Different collection tools tested for the development of new sample collection system for infant feces.

S.N.	Collection tool		Suitability for fecal types
1	Cotton swab		Liquid
2	Stick with propeller		Solid Semi solid
3	Dropper		Liquid
4	Coffee stirrer stick		Solid or semisolid

4.3.2. Bead beating tubes and sample preparation

Two types of bead beating tubes, self-made and readymade commercial tubes, were used during the establishment and validation of high-throughput system for processing fecal samples (Table 4). Self-made tubes were successfully used in the laboratory, but they were not suitable for large scale study because of the time needed for their preparation. Self-made bead beating tubes were prepared by weighing autoclaved 0.25 g of Ø 0.1 mm zirconia/silica beads (BioSpec, Catalog # 11079101z) and 3 Ø 2.5 mm glass beads (BioSpec, Catalog # 11079125) into 2.0 ml screw cap tubes (Sarstedt 72693). Commercial bead beating tubes (pre-filled with

beads) were provided along with MagMAX™ Total Nucleic Acid Isolation Kit, ThermoFisher Scientific (AM1840).

In case of preparation of fecal sample for microbial DNA extraction, for self-made tubes ca. 0.125 g of fecal matter is weighed into the tube (with beads) and 0.5 ml lysis buffer (500 mM Sodium Chloride, 50 mM Tris- Hydrochloric acid pH 8, 50 mM Ethylenediaminetetraacetic acid and 4% Sodium dodecyl sulfate) is added and bead beating is performed. In case of commercial tubes, ~0.25 g feces is weighed into 2 ml Eppendorf tube, homogenized by vortexing with 1X phosphate buffered saline (PBS) pH 7.4 (0.5 ml for infants or 1.0 ml for adult feces samples), then centrifuged at 100 x g for 1 min and 175 µl supernatant is transferred into commercial bead beating tubes with pre-dispensed 235 µl lysis buffer and bead beating is performed.

Table 4 Characteristics of the two types of bead beating tubes tested.

Bead Beating tubes		Properties
Self-made		- Tubes need to be sterilized - Zirconia and glass beads are weighed and filled manually - Capacity: ~1500 µl - Preparation is exhausting and time consuming
Readymade or Commercial		- Sterile tubes with pre-filled beads - Capacity: ~400 µl - Ready to use - Price >10x higher than for self-made

4.3.3. Mechanical cell lysis by bead beating

Mechanical cell lysis can be achieved by two types of bead beating methods, normal bead beating (BB) and repeated bead beating (RBB). Both methods were tested for lysing both adults and infant's fecal samples.

Normal bead beating (BB) is performed at speed of 6.5 m/s (FastPrep®-24, MP Biomedicals), two cycles of 1 min each with 5 min rest between the cycles. Beads and solid particles are pelleted by centrifuging at 16000 x g for 3 min and ~300 µl of lysate is transferred into an

empty Eppendorf tube (for high-throughput system lysates are collected in a 96-deep well plate). Lysates (in tubes or deep well plate) were stored in ice if purification is not done immediately.

Repeated bead beating (RBB) is performed at speed of 5.5 m/s (FastPrep®-24, MP Biomedicals), three cycles of 1 min each with 5 min rest in between the cycles. The lysate is incubated at 95°C for 15 min (VorTemp™ 56, Labnet) with gentle shaking. Beads and solid particles are then pelleted by centrifuging at full speed for 5 min in room temperature and the lysate is transferred into a new 2 ml Eppendorf tube (in 96 deep well plate for high-throughput processing) and stored in ice to minimize the physical damage of DNA. For physical disruption of the remaining intact cells, fresh lysis buffer is added to the tubes again and the whole bead beating step is repeated. Second lysates are transferred into the same tubes or same 96 well deep well plates containing first lysate and mixed properly. Lysates (first and second pooled) were stored in ice if purification was not performed immediately. The amount and form of the sample and volumes of lysis buffer for different type of tubes are shown in Table 5.

Table 5. Amount and form of fecal matter and volume of lysis buffer pipetted into the bead beating tubes for repeated bead beating method (RBB).

Tube type	**Amount and form of fecal matter**	**Lysis buffer (µl)**	
		1st round	**2nd round**
Self-made	1st round: ~0,125 g crude feces + lysis buffer 2nd round: Leftover fecal material + lysis buffer	500	150
Commercial	1st round: 175 µl supernatant + lysis buffer 2nd round: Leftover fecal material + lysis buffer	235	100

4.3.4. Purification of DNA

Lysates obtained after the bead beating step contain released nucleic acids but also all the other components of fecal matter. Purification of nucleic acids was performed with manual method only during the optimization and validation of different methods. Most of the DNA lysates were purified using automatic purification system.

Manual purification method

Precipitation of the extracted nucleic acids (NA) was done as explained in (Salonen et al. 2010). The precipitates were further purified using QIAamp® DNA Mini Kit (250) (Cat. No. 51306) and by following the protocol recommended by the manufacturer.

Automatic purification method

Two different purification systems were used for the optimization and implementation phase. During optimization, KingFisher™ Duo Prime Purification System (Thermo Scientific™) was used which can process 12 samples in one run. After this sytem was scaled up to 96 well format high throughput purification system, Thermo Scientific™ KingFisher™ Flex, which can process 96 cell lysates at once. For both systems, we used MagMAX™ Total nucleic Acid Isolation Kit, AM1840 (Ambion life technologies) according to the recommendation and earlier results by the instrument manufacturer (Thermo Scientific™).

- KingFisher™ Duo Prime Purification System

Samples and reagents were prepared and pipetted in a standard KingFisher™ Deep-well 96 plate following the instructions provided with the DNA extraction kit. Modified MagMAX_Pathogen_High_Vol_Duo protocol was selected for the purification. Purified DNA obtained in elution strip were transferred into new Eppendorf tubes or strips and were stored at +4°C temporarily until concentrations were measured and as required for further experiments, before storing in -20 °C.

- Thermo Scientific™ KingFisher™ Flex purification system

This high throughput system was used to purify up to 96 samples in one run using the standard instructions provided by the manufacturer and available in the kit (Ambion life technologies, MagMAX™ Total nucleic Acid Isolation Kit, AM1840, USA). Thermo Scientific™ Multidrop™ Combi Reagent Dispenser was used to dispense reagents into several 96 well plates required during the run. A detailed protocol for extraction of microbial DNA using this system can be found in appendix B. Modified MagMAX_Pathogen_High_Vol_Duo protocol was used for the preparation of samples and reagents for the purification. Purified DNA samples were eluted in 96 well plate which could be directly stored at +4 °C or -20 °C as specified above.

4.3.5. Preparation and performance validation of wash buffers

Wash buffers for the purification of nucleic acids were self-prepared and tested for their compatibility with the kit and performance. This was motivated by the fact that the volumes of wash buffers required for the automated DNA purification by using KingFisher™ Flex and MagMAX_Pathogen_High_Vol_Duo protocol exceeded those provided in the kit (MagMAX™ Total nucleic Acid Isolation Kit, AM 1840, Lithuania), and we wanted to test whether instead of purchasing expensive extra wash buffers we could use self-made ones. Information on the composition of the commercial wash buffers were not available and hence Literature search was performed to determine the optimum composition of wash buffers for purification of genomic/bacterial DNA (Durst and Staples 1972; Dellaporta, Wood, and Hicks 1983; Piškur and Rupprecht 1995; Yagi et al. 1996a; Somma and Querci 2004; Tan and Yiap 2009; Zumbo 2012; OpenWetWare contributors 2016). Composition of wash buffers prepared in the lab was selected based of the non-hazardous and easy handling property along with excellent performance when compared with commercial buffers. The composition of selected self-made wash buffers (HWI, HWI_E and HWII) are shown in Table 6. Final pH of the wash buffers was maintained near to 8, like that of commercial wash buffers present in MagMAX™ Total nucleic Acid Isolation Kit. For wash buffer 1 with EDTA (HWI_E) pH of the solution (Sodium Acetate and EDTA) was maintained near 8 before autoclaving and ethanol was added afterwards.

Table 6.Composition of self-made wash buffers prepared in laboratory for automated purification of DNA from cell lysates obtained from fecal samples.

Wash Buffer I		Final Concentration
HWI	Sodium Acetate (pH 5.2)	0.3 M
	Ethanol	70%
HWI_E	Sodium Acetate (pH 5.2)	0.3 M
	EDTA (pH ~8.0)	10mM
	Ethanol	70%
Wash Buffer II		
HWII	Tris-HCl (pH 8.0)	10mM
	Ethanol	70%

Microbial DNA from fecal samples were purified using both commercial (wash I and Wash II) and two sets of self-made wash buffers (HWI/HWII and HWI_E/HWII). DNA yields obtained after purification were compared against each other. A qPCR specific to *Bacteroides* bacteria group was performed to test the PCR-compatibility of the DNA extracts obtained with the self-made wash buffers versus the commercial wash buffers by comparing the changes in Threshold Cycle (Cq) along with the variation in template amount, as potential sign of PCR inhibition.

4.3.6. Quantification and quality control of extracted DNA

In case of few samples, NanoDrop™ (ND1000 Spectrophotometer, Thermo Scientific™) and Qubit® Fluorometer (Invitrogen™, CA 92008, USA) were used to measure concentrations of microbial DNA obtained from processed fecal samples. Standard instruction provided by manufacturer for Nanodrop™ and Qubit® dsDNA HS Assay kit were used for measurement of the DNA concentration. PicoGreen® dsDNA quantitation assay was used to measure the DNA concentrations in 96 well layout (leaving 86 samples for measurements in single assay except the standards used to construct standard curve). FLUOstar OPTIMA Microplate Reader (BMG Labtech, Germany) was used to measure the DNA concentrations along with standard instruction provided with Quant-iT™ PicoGreen™ dsDNA Assay Kit Ref. P7589 (ThermoFisher™, USA).

For quality control and visual inspection of the DNA agarose gel electrophoresis was performed. 1 to 3% agarose gels were stained with Midori Green Advanced DNA stain (NIPPON Genetics EUROPE GmbH, Cat no. MG 04), extracted DNA were loaded in the gels to visualize different sizes of DNA e.g. after bead beating experiment and the results were visualized under the UV transilluminator (Gel Doc™ XR+, Molecular imager®, BIO-RAD, USA). Similar procedure was also used to verify the products obtained from PCR.

4.3.7. Establishment of high throughput fecal sample processing system

After all testing and optimizations of tubes, instruments, buffers, purification and quantification methods, an efficient system capable of extracting automatically purified microbial DNA from

96 adults or infant's fecal samples was set up in the laboratory. A detailed protocol for extracting microbial DNA from fecal samples in high throughput fashion using readymade bead beating tubes, FastPrep®-96 and Thermo Scientific™ KingFisher™ Flex purification system is included in appendix B.

4.4. Analysis of fecal bacteria

Extracted fecal DNA samples were subjected to several bacterial genus, species or strain specific PCRs (end-point and real time), and sequencing of 16S rRNA gene amplicons using MiSeq platform to determine the overall microbiota composition.

4.4.1. Full length 16S rRNA PCR

The DNA concentrations obtained from adult samples were in satisfactory level unlike those from swab samples (Esp. for JORVI study samples) which were mostly ≤1 ng/µl and some even in negatives. To ratify the presence of microbial DNA in those low- concentration samples extracted using the RBB method (as described in section 4.3) and purified using self-made wash buffers, PCR reaction was performed to amplify full length of 16S rRNA gene (~1500 bp) using F:5'-GAGAGTTTGATYCTGGCTCAG-3' and R:5'-AAGGAGGTGATCCARCCGCA -3' primers (Wang et al. 2009) obtained from Oligomer Oy, Finland. The 25 µl PCR reaction mixture contained 0.4 µM of forward and reverse primers, 0.2 mM dNTPs (ThermoScientific™), 1.25 U of DNA polymerase (Thermo Scientific™ DreamTaq DNA Polymerase #EP0702), 2.5 µl DreamTaq buffer of same company, 18.75 µl Milli Q water and 1 µl of DNA template. The PCR program contained 2 min. of initial denaturation at 95°C followed by 25 cycles of 20 sec. denaturation at 95°C, 20 sec. annealing at 55°C, 1 min. of elongation at 72°C and finally 5 min. of final elongation at 72°C. The PCR result was visualized using 1% agarose gel.

4.4.2. Real-Time PCR (qPCR) for abundant bacteria

Real-time PCR or qPCR was performed to target the 16S rRNA gene to quantify the dominant bacterial groups among the intestinal microbiota from DNA extracted from fecal samples. Firstly, *Bacteroides* qPCR was performed when testing the self-made wash buffers. Secondly, total bacteria and abundant bacterial groups (*Bacteroides, Clostridium* and *Bifidobacterium*) were quantified from the DNA preps extracted using the high throughput method with self-made wash buffers for several (adult and infant) test samples to evaluate the presence of amplifiable DNA with minimal presence of PCR inhibitors. In the last part of the thesis experiments, a new *L. rhamnosus* GG specific qPCR was developed and validated to replace the old unspecific end-point PCR method. Both PCR methods were used to identify and quantify *L. rhamnosus* GG from the same 3 month's old children samples from FLORA (n=30, control group) and JORVI (n=28, antibiotic and non-antibiotic group) studies and the results were compared. The details of these PCR assays including the bacterial target, primers, template amount, bacterial standards, annealing temperatures and detection temperatures are specified in Table 7.

The qPCR amplification and detection were performed using C1000 Touch® Thermal Cycler (BIO-RAD) with HOT FIREPol® EvaGreen® qPCR Mix Plus (Solis BioDyne, Estonia). All reactions were performed in triplicates and outliers were removed (criteria) prior to further analysis. The qPCR conditions have been described in detail elsewhere (Rinttilä et al. 2004; Salonen et al. 2010). In brief, 25 µl reaction contained 0.20 µM of both primers (Oligomer Oy, Finland), 5 µl of qPCR mix, 14.8 µl of MilliQ water and 5 µl of DNA template. The PCR program included a 1 min initial denaturation at 95°C, followed by 39 cycles of 15 sec denaturation at 95°C, 20 sec annealing (50°C to 68°C) and 30 sec elongation at 72°C, detection from 72°C to 90°C (varies with target bacteria group or specific strain) for 10 to 30 sec. After 39 cycles 1 min of denaturation at 95°C finishing with plate read from 55°C to 95°C with 0.5°C increment in every 0.05 sec for melting curve. Results from qPCR were analyzed by constructing $\log_{10}$ standard curves using known amounts of bacteria present in standards, to convert the threshold cycle (Cq) values into bacterial DNA copy numbers/g of wet feces in as described previously (Salonen et al. 2010).

Table 7. Information about the target bacterial group or specific strain, primers, standards, template amount, annealing temperature, detection temperature and references for qPCR assays used in this thesis.

Target	Primers	Bacterial standard	Template amount (ng)	Annealing temperature (°C)	Detection temperature (°C)	Reference
Total bacteria	F: 5'-TCCTACGGGAGGCAGCAGT-3' R: 5'-GGACTACCAGGGTATCTAATCCTGTT-3'	*Bifidobacterium longum*	0.5	50	82	(Nadkarni et al. 2002)
Bifidobacterium spp.	F: 5'-TCGCGTC(C/T)GGTGTGAAAG-3' R: 5'-CCACATCCAGC(A/G)TCCAC-3'		25	58	85	(Rinttilä et al. 2004)
Bacteroides spp.	F: 5'-GGTGTCGGCTTAAGTGCCAT-3' R: 5'-CGGA(C/T)GTAAGGGCCGTGC-3'	*Bifidobacterium fragilis*	0.5	68	80	
Clostridium coccoides spp.	F: 5'-CGGTACCTGACTAAGAAGC-3' R: 5'-AGTTT(C/T)ATTCTTGCGAACG-3'	*Ruminococcus productus*	0.5	55	85	
***Lactobacillus rhamnosus* GG**	F: 5'-CATTCATGTTGCATCGTCCT-3' R: 5'-GCTTGGCCGGACTAAGTACA-3'	*L. rhamnosus* GG spiked with fecal DNA	12.5	60	72 and 81	Method developed by Pia Rasinkangas and protocol is unpublished

4.4.3. *Lactobacillus rhamnosus* GG specific PCR

End-point PCR assay targeting the Sortase C (srtC) gene (Figure 7) for screening of *L. rhamnosus* GG genome was used to identify and confirm the presence of *L. rhamnosus* GG in fecal samples. It was used to study the prevalence of *L. rhamnosus* GG among 3 month's old children from FLORA (control group) and JORVI study, and co-occurrence of *L. rhamnosus* GG between mother-infant dyads from JORVI study. Sortase C (srtC) gene was amplified using specific forward (5'-AGTGCGACTATTAGCTTTA-3') and reverse (5'-GGATCTTGTGACCTTAATG-3') primers (Table 7) (Oligomer Oy, Finland) and the PCR product (150 bp) was confirmed by visualizing in agarose gel and by sequencing the product using Sanger sequencing. The 25 µl of reaction mixture contained 12.5 µl of master mix (Thermo Scientific™ Phusion High-Fidelity PCR Master Mix, F531), 1.25 µl primers (10 µM), 5 µl water and 5 µl DNA template (1:10 diluted fecal microbial DNA). PCR reaction program included 3 min of initial denaturation at 98°C followed by 30 cycles of 20 sec of denaturation at 98 °C, 30 sec of annealing at 60°C and 1 min of elongation at 72°C, and finally a 5 min of elongation at 72°C.

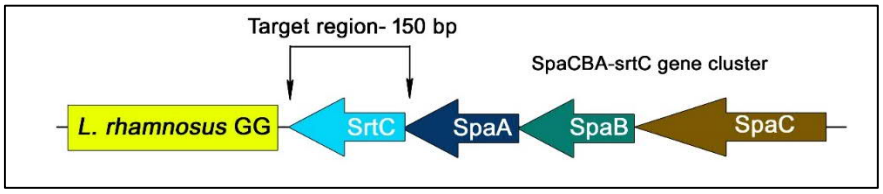


Figure 7. Schematic presentation of the end-point PCR target site (srtC gene) of L. rhamnosus GG, resulting in PCR product of 150 bp.

4.4.3.1. Lactobacillus rhamnosus GG specific qPCR

This method was developed by Pia Rasinkangas, PhD (previously a post-doc researcher in Willem De Vos group, Haartman institute, University of Helsinki and currently Scientist at DuPont, Finland). For improved specificity, a new site was targeted for amplification that contained partial srtC gene and in addition flancking *L. rhamnosus* GG's sequence (srtC

border) resulting in 238 bp product (Figure 8). This method was tested and optimized by myself outside of my thesis project so the details regarding the optimization and validation of this method are not discussed here (manuscript in preparation).

In this project, this newly developed and optimized qPCR assays was used for specific detection and quantification of *L. rhamnosus* GG from fecal DNA extracted from infants at 3 months from FLORA (control group) and JORVI (microbial DNA extracted using the newly developed high throughput system) trials, and the results were compared against the srtC end-point PCR method.

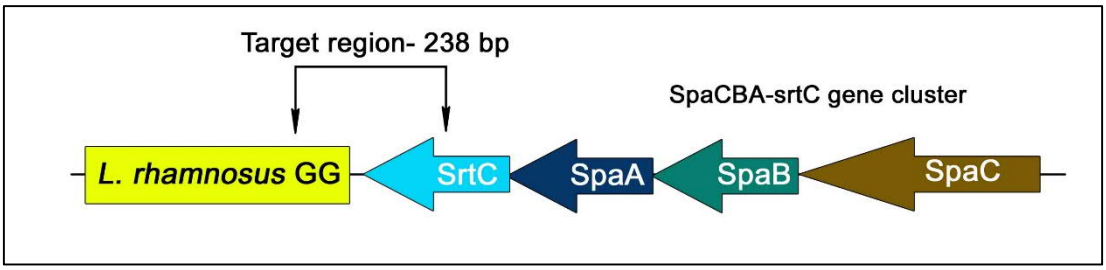


Figure 8. L. rhamnosus GG specific real-time PCR target site of L. rhamnosus GG resulting in PCR product of 238 bp.

For this assay, the entire qPCR protocol was redesigned to ensure specificity of the assay to identify and quantify *L. rhamnosus* GG. Each standard was spiked with 25 ng of fecal DNA (devoid of *L. rhamnosus* GG) to make their amplification conditions analogous to that of the fecal DNA and the template amount for reaction was limited to 12.5 ng. The PCR reaction had denaturation temperature 98°C, 29 cycles and final plate reading from 65°C to 95°C with 0.5°C increment in every 0.05 sec. for melting curve. Results were similarly analyzed as explained under Real Time PCR (qPCR) previously. Detail about primers, template amount, bacterial standard, annealing temperature and detection temperatures are specified in Table 6.

4.4.4. Sequencing

To study the microbiota composition of the selected fecal DNA extracts from infant samples and confirm their suitability for the established library preparation protocol, the samples were analysed using Next generation sequencing. The PCR products from both PCR methods targeting *L. rhamnosus* GG srtC derived from selected fecal samples were sequenced using Sanger sequencing method. Basically, NGS allows us to process large amount of sample at one run and can sequence millions of fragments while sanger sequencing can only handle few samples and produce one forward and reverse reads with 99% accuracy (Gurson 2015). Although NGS have been in use extensively but during planning phase and for confirmation of products before heading towards large scale sanger sequencing is considered as feasible and reliable method. The samples were sequenced at Institute for Molecular Medicine Finland (FIMM), Finland.

Next generation sequencing (Illumina MiSeq)

During optimization phase of high throughput fecal sample processing, microbial DNA extracted from infant's fecal samples using different types of bead beating tubes and methods were sequenced using next generation sequencing method. The samples were prepared by amplifying V3 & V4 regions of the 16S rRNA gene using the locus specific primers 341F and 785Rev (Klindworth et al. 2013) with TruSeq overhangs (F: 5'-CTCGTCGGCAGCGTCAGATGTGTATAAGAGACAG-3' and R: 5'-GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG-3') and barcoding primers according to the protocol kindly provided by Institute for Molecular Medicine Finland (FIMM), Finland that follows the 16S Metagenomic Sequencing Library Preparation (15044223 B) protocol by Illumina® (Amplicon, Clean-Up, and Index 2013) with small modifications. The PCR reaction was performed in 20 µl reaction containing 1 µl primers (0.25µM), 10 µl of master mix (Thermo Scientific™ Phusion High-Fidelity PCR Master Mix, F531), 7µl of MilliQ water and 1 µl of DNA template. Different concentrations and volumes of template DNA were tested and according to the previous unpublished results from de Vos laboratory, 1 µl of template DNA at concentration 1 ng/µl was considered optimum as it provided the highest percentage of PCR-positive samples copared to higher input of fecal DNA. PCR program included an initial denaturation at 98 °C for 30 sec. followed by 25 cycles of denaturation at 98 °C for 30 sec., annealing at 62 °C for 30 sec., extension at 72 °C for 15 sec. and program

ended with final extension at 72 °C for 10 min. The PCR products were visualized in agarose gel as preliminary screening, LabChip® GX Touch system (PerkinElmer®, USA) was introduced for high throughput quantification, quality assessment and capillary electrophoresis separation of purified and amplified DNAs.

The pooled libraries were sequenced with Illumina® MiSeq® platform. Data preprocessing and taxonomic annotation of the reads was performed in R package mare (https://github.com/katrikorpela/mare) by Anne Salonen (my thesis supervisor) before I analyzed them in Excel.

Sanger sequencing

PCR products obtained from old (end-point) and new (real-time) *L. rhamnosus* GG specific PCR methods were sequenced to evaluate the specificity of each method to correctly target *L. rhamnosus* GG from microbial DNA extracted from fecal samples. PCR products from both PCR methods were first purified using GeneJET™ PCR purification kit, Thermo Scientific™, Lot. 00134915 following the protocol provided by manufacturer. Samples for sequencing were prepared by mixing 5 µl of purified DNA (40 ng) with 1.6 ul of 5uM forward primer (used in PCR reaction). NCBI BLAST tool was used to identify the origin of sequences in the PCR products.

5. RESULTS

Results from various experiments are arranged in two sections, section 1 for the optimization and validation of high-throughput system for processing fecal samples and section 2 for studying the prevalence and transfer of *L. rhamnosus* GG in a subset of the extracted samples.

5.1. Section 1: Development and standardization of fecal sample collection system

A high-throughput system is effective only when there is a feasible, efficient and reproducible fecal sample collection technique. Our aim was to develop a new sampling strategy for infant feces to replace the commercial options for fecal home sampling that usually employ ≥15 ml tubes despite of a relatively small sample size, taking a lot of freezer space. Weighing the required amount of sample for DNA extraction takes half of the total processing time and hence eliminating this step from the DNA extraction protocol was our priority. Infant's fecal samples are normally semi solid or watery so they are tricky to collect, and an optimum method for collecting fecal samples of different consistencies was thus required. We tested four collection techniques using real feces of different nature and consistency. The collective results and comments are presented in Table 8.

Table 8. Outcome of the test for four different collection techniques for infant feces.

Collection technique	Fecal type	Sample collected (g)	Remarks
Cotton swab	Solid	0.030	Suitable for collection of watery samples but not possible to collect solid (dry) feces
	Semi solid	0.300	
Stick with propeller	Solid	0.290	Short length of the stick makes it difficult to collect samples properly, useful to collect solid samples only
	Semi solid	0.493	
Plastic dropper	Solid	0.261	Difficult to collect dry samples and unsafe (to cut the tip). Fluffy and watery samples are easier to collect due to suction bulb on top
	Semi solid	0.243	
Coffee stirrer stick	Solid	0.327	Easy to collect all kinds of samples esp. for solid and semi solid samples. For liquid samples, the tube must be placed very close to the collection stick
	Semi solid	0.393	
	Liquid	0.304	

Tests revealed the coffee stirrer stick as the best method for collecting fecal samples and a protocol was made in which study subjects were instructed to collect a pea sized dollop (~250 mg) of baby feces into 2 ml cryo tubes with outer screw grips (less messy than tubes with inward grip), place them in the cool transport container (Sarstedt) and store in -20^{0}C until taken to study center or collected by responsible person (Protocol in Appendix A). By now this protocol has been successfully implemented in large HELMi (https://helmitutkimus.com/) birth cohort study to collect thousands of samples, and the time required for microbial DNA extraction from 96 fecal samples have shorten by 3 hours compared to the old procedure including sample aliquoting as the first step.

5.2. Microbial DNA extraction methods, bead beating tubes and DNA yields

Before scaling up and optimization of DNA extraction in 96-well format, 4 infants and 4 adult's fecal sample were used for initial screening of different extraction methods using two types of bead beating tubes along with two different DNA purification methods. The following methods and components were tested:

DNA extraction methods

a) Repeated Bead Beating (RBB)
b) Bead Beating (BB)

Bead beating tubes

a) Self-made bead beating tubes (ST)
b) Commercial (ready-made) bead beating tubes (RT)

DNA purification systems

a) QIAamp® DNA Mini Kit (250) – Manual purification
b) KingFisher™ Duo Prime purification system – Automatic purification

The manual DNA extraction method that had been previously validated in our laboratory for adult fecal samples (Salonen et al. 2010) was used a basis for upscaling the DNA extraction method; employing the option a) in the above listing. Before this project, the initial validation

of for the use of KingFisher™ Flex purification system for semi-automated 96-format fecal DNA extraction had been performed, indicating that for adult samples the DNA yields were comparable between the manual and automated DNA purification method (data not shown). However, as the bead beating procedure from the manual system is not feasible for 96-format, we needed to adjust and optimize the method for the upscaled protocol. Hence, we compared the performance of self-made bead beating tubes, adapted from the manual protocol to the commercial bead beating tubes that come as part of the DNA extraction kit (MagMax™ Total nucleic acid Isolation Kit, AM1840) that had been selected for automated DNA purification based on the recommendation of the KingFisher™ technical support. As the two types of bead beating tubes arguably differ on their sharing efficiency due to differences in tube volume and shape (Table 4 and 5), we also tested if longer (repeated beat beating, RBB) or shorter (simple beat beating, BB) treatment yields better results with the commercial bead beating tubes that are needed for the 96-format. Manual extraction was used as a reference for the automated purification with the KingFisher™ Duo Prime purification system, a 12-sample set-up version of the KingFisher™ Flex Purification System operating in 96-format. Four infant (aged 1 to 5 months) and four adult fecal samples were used for these tests. Due to the small sample size the results are descriptive and no statistical testing was performed.

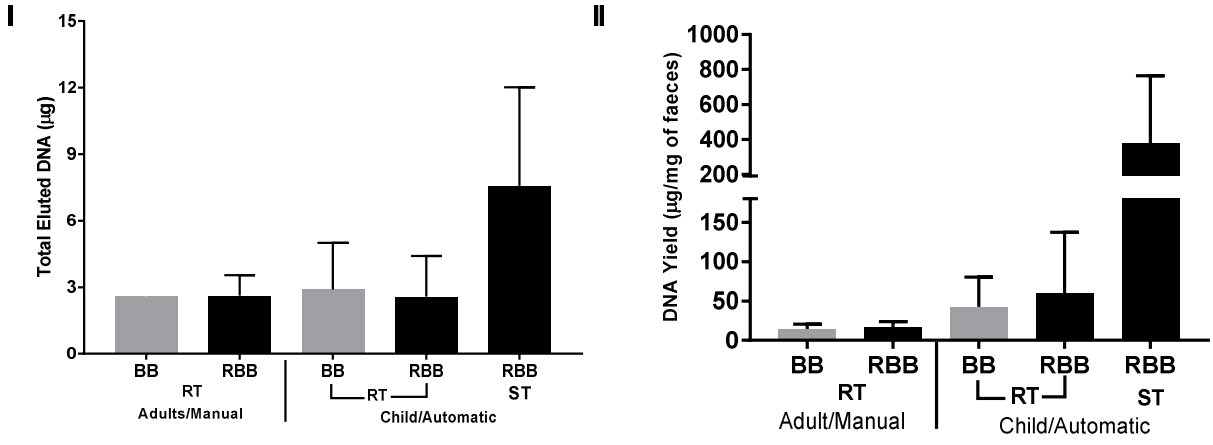


Figure 9. DNA yields in absolute amount (I) and normalized per sample weight (II) from different samples and treatments. Adult and baby fecal samples were processed using Repeated Bead Beating (RBB) and Bead Beating (BB) methods in readymade (RT) and self-made bead beating tubes (ST) using manual and automatic DNA purification systems.

Based on the fluorometric measurements, the largest difference in the DNA yields derived from the tubes; the self-made bead beating tubes gave ca. 2.7-fold higher absolute DNA yield (7.57 µg) when normalized to the sample weight compared to the readymade tubes in case of child's fecal samples (Figure 9). This is logical given that the volume of readymade tubes is 2.5x smaller (Table 5). In the readymade tubes that are needed for high throughput system, the average DNA yields were comparable between the BB and RBB methods for adult samples both in absolute (2.58 and 2.62 µg) and normalized (14.58 and 16.17 µg/mg feces) terms. Also in case of the infant samples, the yields between the BB method (2.88 µg) and RBB method (2.57 µg) were quite similar, but the normalized DNA yields were lower with the BB method (42.55 vs 59.70 µg/mg). An agarose gel was additionally used to determine the effectiveness of bead beating methods and tube types (Figure 10).

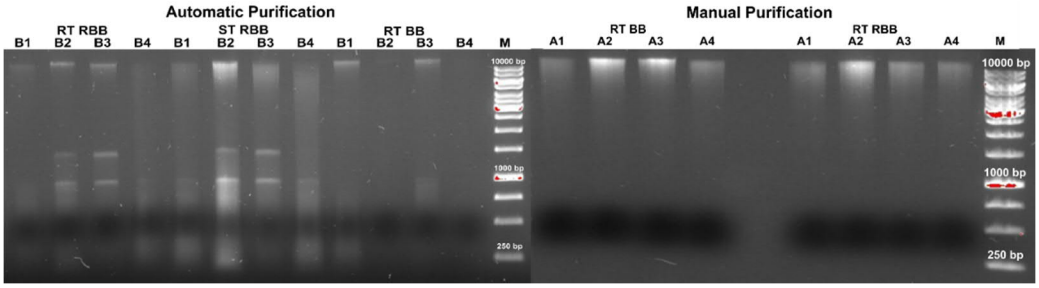


Figure 10. Agarose gel run of the genomic DNA extracted from infant (B1 – B4) and adult (A1 – A4) fecal samples using two types of bead beating tubes, commercial (RT) and self-made (ST), bead beating methods (RBB and BB) and purification methods (automatic and manual).

In line with the fluoreometry-based DNA quantification, the presence of smears throughout the samples extracted using self-made tubes indicated higher amount of DNA present in the eluates. Adult fecal samples gave similar bands (>8000 bp) from both DNA extraction methods (BB and RBB). For the infant samples, DNA extraction with commercial tubes gave 3 times less DNA (µg) and 8 times less DNA per mg feces than self-made tubes (Figure 9 II). Similarly, DNA bands and smears observed in the gel also indicate higher amount of DNA when using self-made tubes and RBB method for infant samples. With readymade tubes and RBB method, the banding pattern (in agarose gel, Figure 10) was comparable to that of self-made tubes,

suggesting that the readymade bead beating tubes can provide sample shearing efficiency that is comparable to that of self-made tubes.

5.3. Amplification of abundant bacteria using qPCR

To estimate the amount of amplifiable bacterial DNA in the fecal DNA extracts, 16S rRNA gene of the most abundant bacterial groups (*Bacteroidetes, Clostridium* and *Bifidobacteria*) and total bacteria were targeted using qPCR.

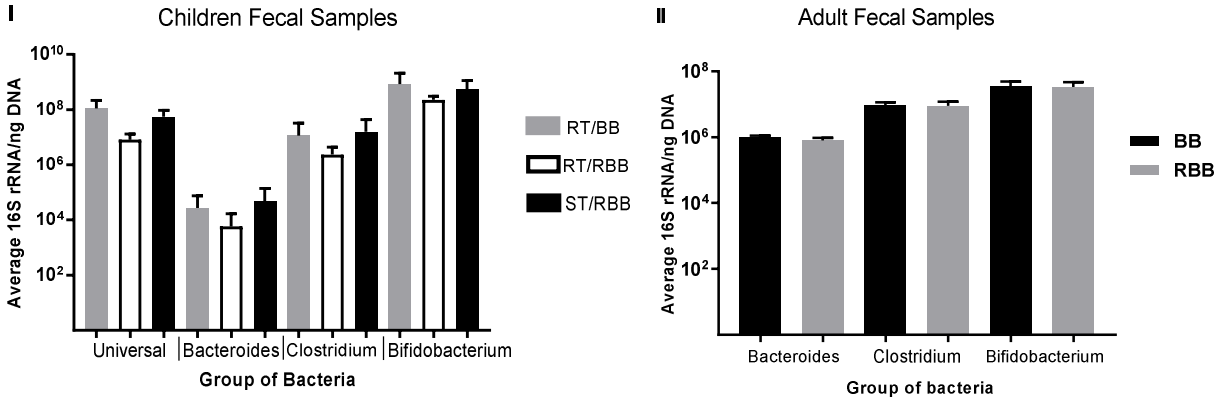


Figure 11 qPCR results for the abundant bacteria in (I) infant and (II) adult fecal DNA extracted using Bead Beating (BB) and Repeated Bead Beating (RBB) method with commercial (RT) and self-made (ST) bead beating tubes. Adult fecal samples were extracted using readymade tubes and purified manually.

For the total bacteria and all tested bacterial groups in infant samples purified with automated system, BB method gave roughly 0.3 log higher counts than the RBB method when using the commercial tubes, while using the self-made tubes the counts were comparable to the BB method (Figure 11 I). Amplification on manually purified adult's fecal DNA (Figure 11 II) extracted using readymade tubes and bead beating methods (BB and RBB) produced similar amplification (~10^6 to ~10^7) for all the bacterial groups.

5.4. 16S rRNA gene amplicon sequencing

Amplicons of the 16S rRNA gene (V3-V4 region) were sequenced with Illumina MiSeq platform using infant's fecal DNA extracted with different bead beating tubes and methods as explained above to study their potential effects on the bacterial composition. The relative proportions of the bacterial phyla (Figure 12 I) and the ratio of gram negative and gram-positive bacteria (Figure 12 II) were found to be very similar between the methods, individuality driving the between-sample (B1-B4) differences.

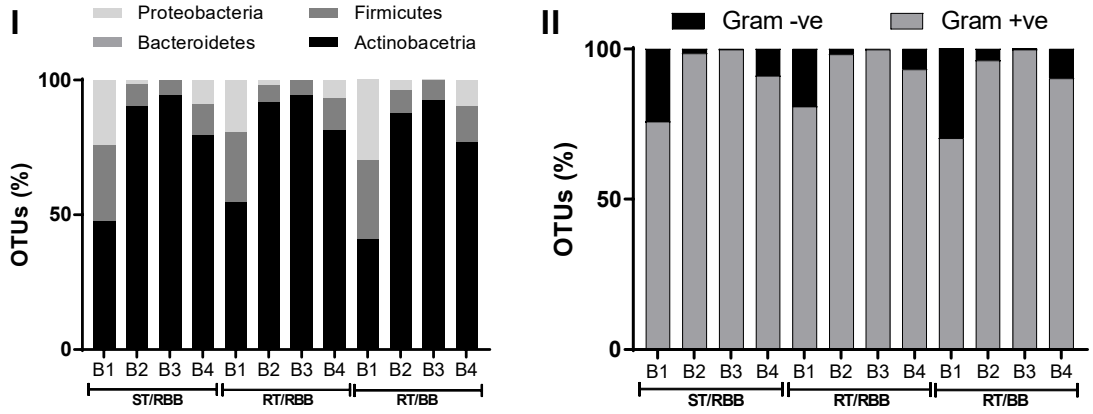


Figure 12 Gram positive or negative abundant bacterial phyla obtained from infant's fecal DNA (B1-B4) extracted using two types of bead beating tubes (self-made and commercial) and two extraction methods (BB and RBB)

In all samples *Actinobacteria*, more specifically *Bifidobacteria*, dominated the communitiy with relative abundance from 50% to 90%, *Proteobacteria* and adult-type *Bacteroidetes* and Clostridia being less abundant. Genus-level analysis indicated that both DNA extraction methods (BB and RBB) and bead beating tubes (RT and ST) successfully extracted DNA from both gram positive (*Bifidobacterium, Anaerostipes and Streptococcus*) and gram-negative bacteria. In samples B1, abundancy of *Proteobacteria* was ~25% which together with Bacteroidetes is a major phylum of gram negative bacteria within the gut microbiota. As gram positive bacteria, containing thicker cell wall, typically account for more than 50% of the bacteria in fecal samples, a successful extraction of DNA from these bacteria testifies the effectiveness of used DNA extraction methods. In this study, the samples extracted with commercial versus self-made bead beating tubes gave similar, high abundance of gram positive bacteria (Figure 12 II).

5.5. Commercial wash buffers vs. self-made wash buffers for purification of DNA from fecal extracts

Commercial wash buffers (MagMax™ Total Nucleic Acid Isolation kit) used for purification of fecal DNA extracts turned out to be insufficient for fecal samples, raising a need to develop affordable self-made wash buffers. Newly composed self-made wash buffers were tested against MagMax's commercial wash buffers for their performance in high-throughput purification (Thermo Scientific™ KingFisher™ Flex Purification System) system. Two self-made wash buffers were prepared and tested to examine their compatibility of the system and impact on the DNA yields and quality (Table 9).

Table 9 Different type of wash buffers tested for purification of fecal DNA extracts

Commercial wash buffer	**MagMax™**	
Self-made wash buffers	Wash Buffer I	**HWI**
		HWI_E
	Wash Buffer II	**HWII**

Microbial DNA from fecal samples were extracted from adult and child fecal samples using RBB method with commercial bead beating tubes. Self-made and commercial wash buffers were used for purification of the same infant samples (B1-B4) as used in previous experiments for testing of high-throughput system. The performance of the wash buffers was evaluated by comparing the change in Threshold Cycle (Cq) against the template amount used in PCR. In the absence of PCR inhibition, Cq decreases linearly as a function of increased amount of template DNA (Figure 13).

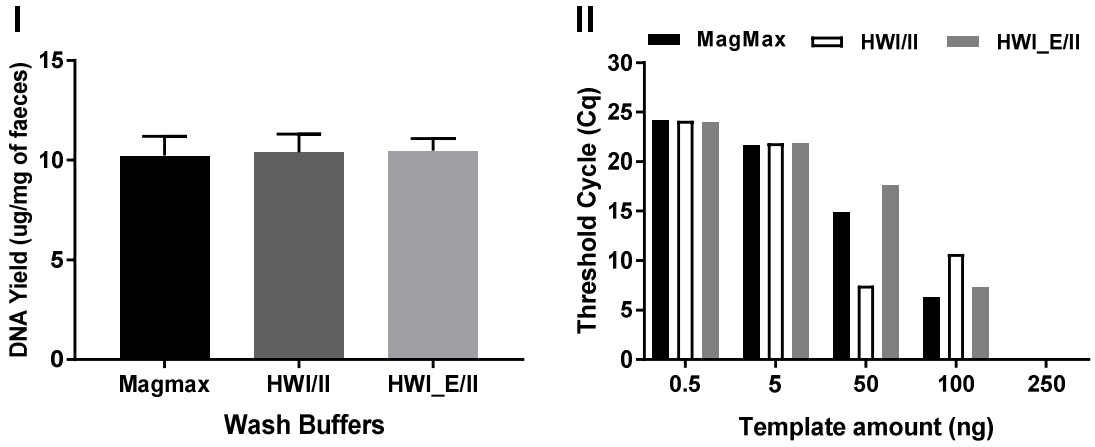


Figure 13 DNA yields obtained from extraction using different wash buffers for purification (I) and comparison of Cq values against template amount used in qPCR (II)

The difference between DNA yields (Fig. 13 I) were minute (10.24, 10.39 and 10.48 µg/mg of feces) among different wash buffer sets. Comparison of the Threshold Cycle (Cq) values in *Bacteroidetes* qPCR showed no differences between the buffers with low template amounts (0.5 and 5 ng). With higher input template (50 to 100 ng), some differences emerged (Fig 13 II). The EDTA-containing self-made buffer (HWI_E/II) and MagMax™ buffer showed an expected, decreasing trend of Cqs, while the other self-made buffer set HWI/II gave inconsistent results and no amplification was recorded for any sample when using the highest amount of template DNA (250 ng). We concluded that at concentrations that are used for preparation of the sequencing libraries (1-5 ng), HWI_E/II (first wash buffer with EDTA) buffer set provided similar or better amplifications than the MagMax™ commercial wash buffer and was hence implemented in the final protocol for purification of fecal DNA extracts in 96 well format.

5.6. High-throughput system for processing fecal samples

From the above-described experiments and optimizations, we established a method for the automated DNA extraction of fecal samples based on RBB method for cell lysis, KingFisher™ Flex Purification system and MagMax™ Total Nucleic Acid Isolation Kit along with self-made wash buffers (HWI/II and HWI_E/II). To enable bead beating in a high-throughput manner, FastPrep®-24 homogenizer was replaced by FastPrep®-96 instrument. Following is the setup for 96 well format high-throughput microbial DNA extraction and purification system.

DNA extraction/purification Kit: MagMax™ Total Nucleic Acid Isolation kit

DNA extraction method: Repeated Bead Beating (RBB)

Grinding, lysis and homogenization: FastPrep®-96

Incubation: VorTemp™ 56 microplate shaking incubator (56 tubes per run)

Wash buffers: Self-made wash buffer I with EDTA (**HWI_E**) and wash buffer II (**HWII**)

Reagent dispenser: Thermo Scientific™ Multidrop™ Combi Reagent Dispenser

DNA purification instrument: Thermo Scientific™ KingFisher™ Flex

DNA quantification/storage: Quant-iT™ High-Sensitivity dsDNA Assay Kit/ 96 well plates

5.7. Application of high-throughput microbial DNA extraction and purification system

5.7.1. Extraction of microbial DNA from paired infant and mother fecal samples

Fecal samples collected in the context of the JORVI study were extracted using the newly developed and optimized high-throughput system. DNA was extracted from 60 mother-infant pairs sampled longitudinally, the number of available samples at the time of this project totaling to 647 samples: 342 from the mothers (mean age 33.5 years) and 292 from the infants (from the first day until 3 months) excluding dropout samples. Of the samples, 88% had been

collected using spatula and were easy to weigh for the extraction process and the remaining 12% had been collected using swabs and were directly soaked into PBS.

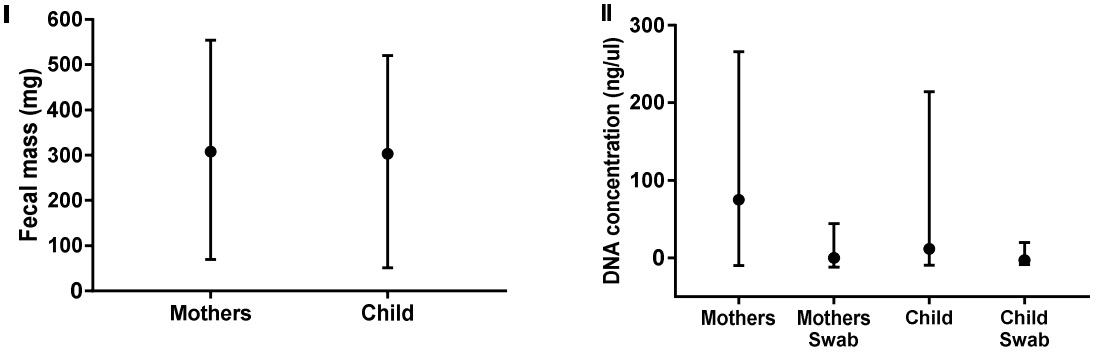


Figure 14 (I) Amount of fecal material used to extract DNA from mothers and children samples (II) amount of DNA extracted from different sample types (feces and swabs)

250 mg is the optimum amount of starting material (feces) required to extract DNA but depending on sample type (solid, semi-solid and liquid), the amount of starting material varied. For adult starting material varied from 69 mg to 554 mg (semi solid feces containing fibers and undigested foods) and in infants it varied from 51 mg to 520 mg (Figure 14 I). The mean DNA concentrations for fecal samples from adults and infants were 74.88 ng/µl and 11.66 ng/µl, respectively (Figure 14 II). The amount of fecal material collected using swab was evidently very much lower than in the normal samples as most swabs had hardly any visible fecal matter. The mean concentrations from mother (adults) and infant's swab samples were less than 1 ng/µl (mothers: 0.058 ng/µl and infants: -2.94 ng/µl), negative concentrations are observed due to the calculation of concentration based on amount of fluorescence detected. Due to the scarcity of DNA in the swab samples, full length 16S rRNA gene PCR was performed on randomly selected samples varying in concentrations to confirm the suitability of the selected method for DNA-based bacterial analysis.

5.7.2. 16S rRNA full length PCR for adults and infant's fecal DNA samples

Out of 342 adult samples, 50 were collected using swabs and 31 had negative DNA concentrations. In case of 292 infants, 267 were collected normally and 25 were collected using swabs, amongst them ~40% (21 swabs and 97 normally collected) had negative DNA concentrations. To confirm the presence of bacterial DNA in these eluates, several samples with negative and positive concentrations were randomly selected and 16S rRNA gene full length PCR was performed.

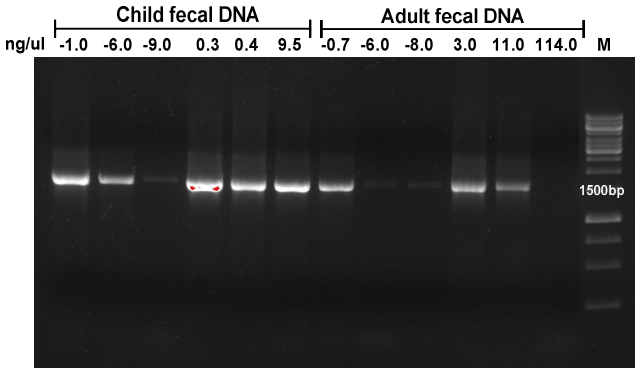


Figure 15 Agarose gel result of 16S rRNA gene full length PCR performed on adults and children fecal DNA with different concentrations (ng/µl), extracted from JORVI study samples.

Among the tested samples, all samples gave PCR products (~1500 bp) except one adult sample which had the highest concentration of 114 ng/µl (Figure 15). The lowest negative concentrations (for example -6.0, -8.0 and -9.0) gave the weakest bands. All DNA extracted from children's fecal samples gave strong bands unlike one adult's DNA with highest concentration (114.0 ng/µl). Overall, strong and clear bands from all the samples enabled us to confirm the presence of amplifiable microbial DNA in the extracted samples including those with negative concentrations. Loading very high amount of template DNA in PCR might limit reagents used in reaction (esp. magnesium ions to stabilize the DNA structure) and produced less or no amplicons.

5.8. Section 2: prevalence and abundance of *Lactobacillus rhamnosus* GG in a subset of the extracted samples

This section will provide the information about the results obtained from experiments performed towards the following objectives.

a) To understand the prevalence of *L. rhamnosus* GG among 3 month's old babies from two studies done ~10 years apart, placebo group of FLORA (n=30) and JORVI (antibiotic and non-antibiotic group) study (n=28).

b) To study the co-occurrence of *L. rhamnosus* GG in mothers and their infants from birth until 3 months after giving birth.

Note: For tables and figures *Lactobacillus rhamnosus* GG (*L. rhamnosus* GG) will be abbreviated as LGG.

5.9. Prevalence of *L. rhamnosus* GG among 3-month-old infants

The prevalence of *L. rhamnosus* GG among 3 months old infants from FLORA study (placebo group) and JORVI study (antibiotic and non-antibiotic group) was studied using PCR method which targeted *L. rhamnosus* GG specific sortase C gene (Figure 16).

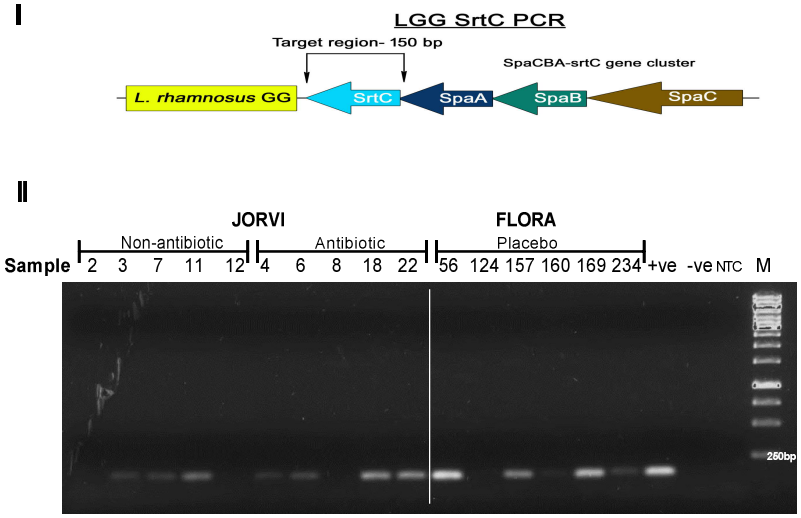


Figure 16 (I) Target region (gene) of sortase C gene present in L. rhamnosus GG (II) Agarose gel electrophoresis to study the presence of sortase C gene as marker of L. rhamnosus GG *in 3 month's old babies' fecal DNA extracted from JORVI and FLORA studies. Sample ID A33 from active group (infants who took probiotic capsules containing LGG) of FLORA study was included in reaction as positive (+ve) control and fecal DNA containing* LGG under detection limit *as negative (-ve) control including NTC (non-template control).*

In total 58 samples were analyzed and 35 (60.3%) of them were positive for *L. rhamnosus* GG. A subset of samples is visualized in Figure 16 II, showing that samples 3, 7, 11, 4, 6, 18, 22, 56, 157, 160, 169 and 234 gave positive bands (150 bp) for *L. rhamnosus* GG.

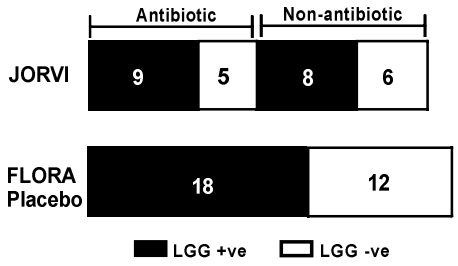


Figure 17 Prevalence of L. rhamnosus GG among 3 month's old children from JORVI and FLORA studies

Overall ca. 60 % of the studied 3 month's old infants were positive for *L. rhamnosus* GG in both JORVI study (n=28) and FLORA study (n=30). In JORVI study, number of *L. rhamnosus* GG positive in both antibiotic and non-antibiotic group were 9 and 8 respectively (Figure 17). Only infants positive for *L. rhamnosus* GG at 3 month's (3m) age were selected for further analysis and *L. rhamnosus* GG srtC PCR were performed on the fecal DNA extracts collected at earlier time points, i.e. at 1 month (1m), 2 weeks (2w), 1 week (1w) and at birth (B).

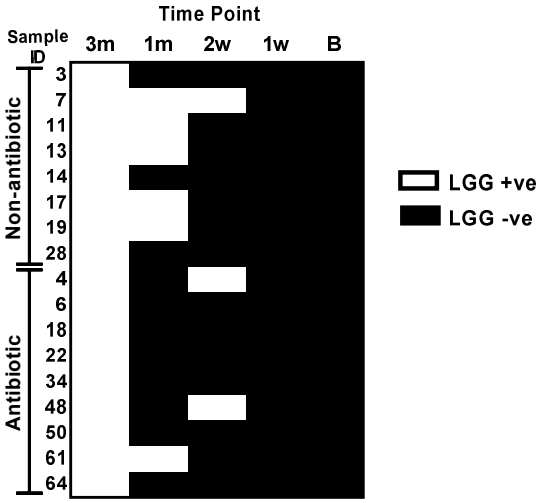


Figure 18 Prevalence of L. rhamnosus GG (LGG) in infants at different time points after birth (B) until 3 months when all the selected 17 infants were positive for LGG.

All infants were negative for *L. rhamnosus* GG at the time of birth and after one week (Figure 18). At the age of 2 weeks, 3 infants and 6 infants at the age of 1 month were *L. rhamnosus* GG positive. Only 1 infant was positive at the age of 1 month from antibiotic group compared

to the 5 *L. rhamnosus* GG positives from non-antibiotic group. Detailed information about the use of different antibiotics by mothers and prevalence of *L. rhamnosus* GG at birth and 3 months after birth is listed in Table 10.

5.10. Co-occurrence of *L. rhamnosus* GG in mother-infant pairs

To study the potential transfer of *L. rhamnosus* GG from the mother to her child, we utilized the fact that paired mother-infant samples were available from the same time points. Only children positive for *L. rhamnosus* GG at 3 months were selected for this analysis to study the early temporal patterns. Hence, we analyzed mother's fecal DNA from all samples (n=28) and from all-time points, i.e. birth (B), 2 weeks (2w), 1 month (1m) and 3 months (3m) using the same srtC PCR method as above. Comparison between the prevalence of *L. rhamnosus* GG amongst the mothers and their infant's samples collected at same time points was used to speculate if infants get these specific bacteria from their mother.

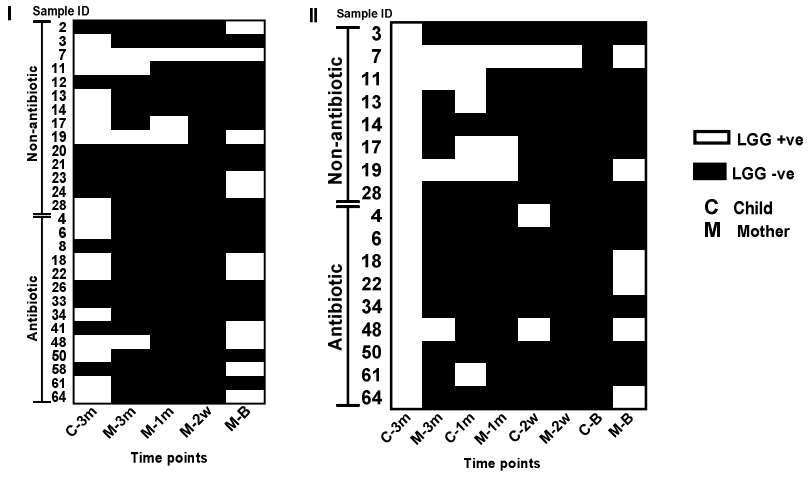


Figure 19 (I) Positivity of srtC PCR as a marker of L. rhamnosus GG (LGG) carriage in 3-month-old children (C; N=28) and their mother's (M) fecal DNA samples at different time points, from birth (B) to 3 months (3m) (II) Comparison between all the time points for presence of LGG in mothers and their children who were all positive for LGG at the age of 3 month (N=17).

Circa ~60% of the infants (n=28) were positive for *L. rhamnosus* GG at 3m whilst only ~14% of the mothers (n=28) were positive at the same time point, and ~39% when they gave birth

(Figure 19 I). Among the mothers of infants that were all positive at the age of 3m (n=17), ~35% were positive for *L. rhamnosus* GG at the time of childbirth. In other words, only 4 mothers and their child were both positive at 3m time point and 13 mothers were negative for *L. rhamnosus* GG while their child was positive. Only one mother (ID# 7) was positive at all-time points and her child was also positive at all-time points (Figure 19 II) except at birth (B). Status of occurrence of *L. rhamnosus* GG among mother-infant pairs and antibiotics used by pregnant mothers can be observed in Table 10.

Table 10 A detailed information about the use of intravenous antibiotics by pregnant mothers during parturition (n=28) and antibiotics used by infants within 3 months after birth in JORVI study, and occurrence of LGG at the time of birth and at 3 month for mother-infant pairs.

		Mothers			Infants		
	ID	**LGG +ve (at birth)**	**Antibiotics**	**LGG +ve (3m)**	**LGG +ve (at birth)**	**Antibiotics (≤3m)**	**LGG +ve (≤3m)**
ANTIBIOTICS	**04**	0	***Pen**	0	0		1
	06	0	**Cef, Met**	0	0	**Amo**	1
	08	0	***Pen, *Cep**	0	0		0
	18	1	***Pen, *Cep**	0	0		1
	22	1	***Pen**	0	0		1
	26	0	***Cef**	0	0		0
	33	0	***Cef, *Cep**	0	0		0
	34	0	***Pen**	0	0		1
	41	1	***Cef**	0	0		0
	48	1	**Cef**	1	0		1
	50	0	***Cef**	0	0		1
	58	1	***Unk**	0	0		0
	61	0	***Cef, *Cli**	0	0		1
	64	1	**Cef, Met, Cep**	0	0		1
Total		**6**		**1**	**0**		**9**
NON-ANTIBIOTICS	**02**	1		0	0		0
	03	0		0	0		1
	07	1	**Cef, Met**	1	0		1
	11	0		1	0		1
	12	0		0	0		0
	13	0		0	0		1
	14	0	**Amo**	0	0		1
	17	0	**Cep**	0	0		1
	19	1	**Cep**	1	0	**Amo**	1
	20	0		0	0		0
	21	0		0	0		0
	23	1		0	0		0
	24	1		0	0		0
	28	0		0	0		1
Total		**5**		3	0		8

Amo- Amoxicillin, Cef- Cefuroxime, Cep- Cephalexin, Cli- Clindamycin, Met- Metronidazole, Pen- Penicillin and Unk- Unknown

* Pregnant mothers completed at least one dosage of antibiotics before birth

5.11. *L. rhamnosus* GG srtC border qPCR

During the project, it turned out that the sortase C gene targeted in the above experiments using srtC PCR for identification of *L. rhamnosus* GG was not fully specific for *L. rhamnosus* GG and hence amplification of DNA also from other bacteria possessing the same sortase C gene cannot be excluded. Presence of sortase C gene in different species of lactic acid bacteria (*Lactobacillus casei, Lactobacillus paracasei* and *Lactobacillus rhamnosus*) have questioned the previous detection method for identification of specific *Lactobacillus rhamnosus* GG strain. For accurate identification of *L. rhamnosus* GG present in fecal DNA extracts, this new qPCR assay (srtC border) targets the border of *L. rhamnosus* GG specific gene and partial sortase C gene (Figure 20).

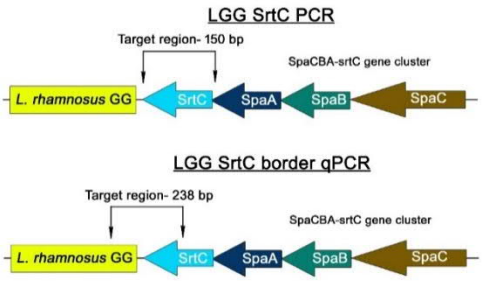


Figure 20 Difference between target regions used in PCR and qPCR for identification and quantification of L. rhamnosus GG.

Vigorous testing and validations were performed for optimization of new *L. rhamnosus* GG detection and quantification method. Fresh *L. rhamnosus* GG standards were prepared from genomic DNA (standard *L. rhamnosus* GG count: 6.12E+05 cfu/µl) extracted from pure cultures in De Man, Rogosa and Sharpe (MRS) broth and plates with manual grid counting of colonies. Six serially diluted standards were optimized and used to identify and quantify *L. rhamnosus* GG from microbial DNA extracted from children fecal samples collected during JORVI (n=28) and placebo group of FLORA (n=30) studies.

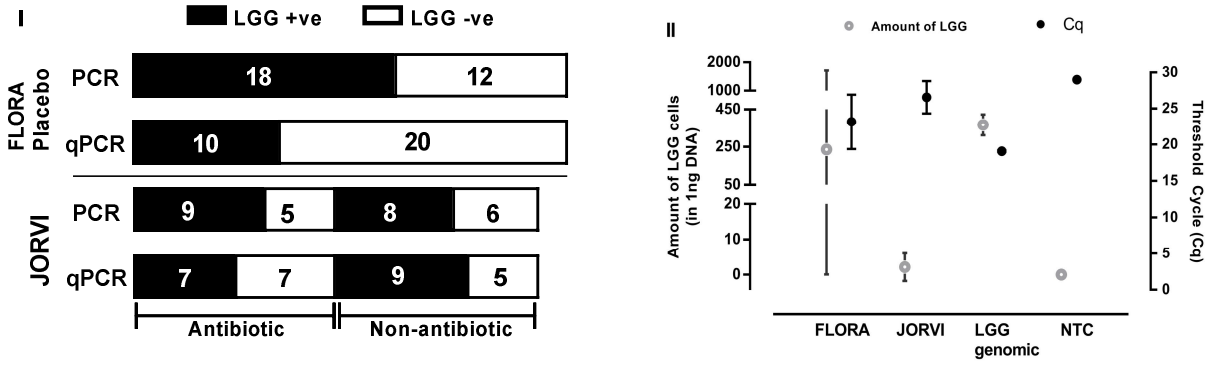


Figure 21 (I) Comparison between the number of L. rhamnosus GG positive samples between the old assay (LGG srtC-PCR) and new assay (LGG srtC border qPCR) on selected 3 months old children samples from JORVI and FLORA study (II) Scattered plot showing the average count (with standard deviation) of LGG detected with the new qPCR assay (LGG srtC border) against threshold cycle (Cq) value on samples from 3 month old children in FLORA (n=30) and JORVI (n=28) study, along with standard LGG genomic DNA and water as non-template control (NTC.)

Among FLORA samples (n=30), ~60% tested positive using the old PCR method, while only ~33% were positive using the new srtC border qPCR method (Figure 21 I). In other words, the number of positive samples decreased by ~50%, indicating high ratio of false positive results with the old PCR assay. There were no significant differences in positives among JORVI samples (n=28), 2 children in antibiotic group who were positive when tested with the old method were negative with qPCR. Only 58 samples were analyzed with this new method excluding all other time point samples from mothers and their child, thus we did not get a complete analyzed result like from end point PCR method, due to which we are limited to use the non-reliable result from old unspecific method.

The new qPCR method enabled also quantification of *L. rhamnosus* GG in the samples. The quantity of *L. rhamnosus* GG detected from different study groups varied significantly. The average number of bacteria (calculated per 1 ng template DNA) detected from FLORA, JORVI, LGG genomic DNA and NTC were 236, 2, 367 and 0 respectively (Figure 21 II). The copy numbers ranged from 0.02 to 1715.4, indicating that this highly sensitive method can detect up to one *L. rhamnosus* GG cell present in the fecal DNA extract. The Cq values greater than 29.0 were considered as negative (as in NTC) since no *L. rhamnosus* GG were detected after 29th

qPCR cycle. Pure *L. rhamnosus* GG's genomic DNA crossed the threshold after 18th cycle. Few amplified samples, using PCR and qPCR, were randomly selected and sent to Sanger sequencing. The sequences derived from the old assay gave multiple BLAST hits in contrast to the limited hits with the new assay (Table 11) as expected.

Table 11 Primer details and BLAST hits obtained for sequences of PCR amplicons generated with the two *L. rhamnosus* targeting assays specified in the text.

Assay	**LGG srtC PCR**		**LGG srtC_border qPCR**	
Primers	Fwd- AGTGCGACTATTAGCTTTA		Fwd- CATTCATGTTGCATCGTCCT	
	Rev- GGATCTTGTGACCTTAATG		Rev- GCTTGGCCGGACTAAGTACA	
Hits	**Bacteria**	**Strain**	**Bacteria**	**Strain**
	Lactobacillus paracasei	IIA	*Lactobacillus casei*	LOCK919 Plasmid pLOCK919
		KL1	***Lactobacillus rhamnosus***	**WQ2**
		L9		**GG, ATCC 53103**
		CAUH35		
		N1115		
		S.sp. paracasei 8700:2		
		S.sp. paracasei JCM 8130		
		ATCC 334		
	Lactobacillus casei	Zhang		
		12A		
		W56		
		BD-II		
		LC2W		
		BL23		
		LOCK919		
		Plasmid pLOCK919		
	Lactobacillus rhamnosus	**WQ2**		
		GG, ATCC 53103		

The old assay (LGG srtC PCR) targets only sortase C gene (srtC) and in addition to *L. rhamnosus* GG had also detected multiple strains of *L. casei* and *L. paracasei*, which makes this assay unspecific. The new qPCR method was specific to two strains of *L. rhamnosus* (WQ2 and GG) and a 29,768 bp plasmid pLOCK919 from *L. casei.* This results indicate that the newly tested qPCR method substantiated the specific detection and quantification of LGG from child and adult's fecal DNA extracts.

6. DISCUSSION

6.1. Development of a fecal sample collection system that circumvents the need for sample aliquoting prior to DNA extraction

Collection of fecal samples and their appropriate storage until further processing is a critical step in conducting human microbiome studies. Maintaining the natural integrity of sample is a must to avoid biased microbiota results, where slight interference can question the quality of the outcome (Choo, Leong, and Rogers 2015). In this study, among the four different collection methods tested for infant fecal samples, coffee stirrers became the best option simply due to length of the stick, easiness in handling and convenience in collection of fecal scoops except for watery samples, which need to be collected with cotton swabs.

No standard quantity of fecal matter have been defined for DNA extraction, the amount ranges from 25 to 200 mg depending upon extraction procedures (Mirsepasi et al. 2014; Wesolowska-Andersen et al. 2014; Hart et al. 2015). Our protocol instructs parents to collect a minimum of ~250 mg of infant feces, which is required for the modified RBB DNA extraction method, and hence provides a "ready-to-use" sample for DNA extraction without the need for aliquoting the samples. The entire collection tube is weighed to record the sample weight for further calculations, in this manner the time it takes to thaw samples after -80°C storage can be utilized for weighing. Then the entire sample is homogenized with PBS, without a backup for crude feces, which is a downside of this system. Hence, duplicate sampling is recommended if crude fecal matter is needed after DNA extraction from the samples. The current technique suits well to a high throughput system because the labour intensive sample weighing step is eliminated and hence the sample preparation time could be reduced as much as by ~50% while maintaining the required cold chain throughout the process, from the sample collection until DNA extraction.

6.2. Optimization of key parameters for high throughput fecal DNA extraction: Bacterial cell disruption by bead beating and purification and quantification of DNA

An optimum fecal DNA extraction system should efficiently deliver high quality DNA that accurately represents the original bacterial community. The system is considered high throughput if it can process large number of samples at once, typically in a single run using semi-automated or fully automatic systems. In our newly validated and optimized high throughput fecal sample processing system, thawed fecal samples undergo mechanical disruption for release of the bacterial DNA in the presence of lysis buffer. In this work, we studied the effect of bead beating tubes and protocols on the fecal DNA yield.

We first studied how the performance of the self-made tubes (ST) that had a long history of use in the laboratory as part of the manual method compare to that of the readymade tubes (RT) that are needed for the high throughput DNA extraction. With ST, the DNA yield was over 6-fold higher (376.25 µg/mg vs 59.707 µg/mg) from children's samples. The major difference between RT and ST is the volume of the tubes and the starting amount of fecal material used in the extraction. In ST it is ~125 mg and in RT only 14-20% (175 µl) of homogenized fecal suspension (~250 mg of feces in 500 µl PBS for infants and 1000 µl for adult samples) fits in the tube. Despite the better performance of the STs, they were not a realistic option for high throughput DNA extraction system as manual dispension of the beads for hundreds and thousands of samples is not feasible. As we were unable to find RT with larger volume than the ones we tested, and their DNA yield was sufficient for the downstream analyses, they were selected for further use. However, any comparisons to be made between the DNA yields of the manual and automated DNA extraction protocol in question should acknowledge the strong influence of the differently sized bead beating tubes as part of the protocol.

Our tests indicated that the readymade tubes as part of the selected DNA extraction kit were suitable for extraction of microbial DNA from adult and baby fecal samples. Despite the decreased DNA quantity (lower DNA yield and bacterial counts in qPCR), the readymade tubes proved to provide fecal DNA that qualitatively compared well to the manual system; they were efficient in breaking down also the hard-to-lyse gram positive bacterial cells and produced a representative DNA profile based on the sequencing results. As the bead beating methods (RBB vs. BB) did not yield observable differences in the tested infant samples, we decided to

stick, for consistency, to the RBB protocol that we have previously shown to be superior for adult samples in terms of the species representativeness in fecal microbiota profiles (Salonen et al. 2010).

Using prefilled bead tubes reduces the extra effort and time required to weigh and manually fill tubes with autoclaved beads. Along with the tubes, KingFisher's automatic purification method drastically reduced the hands-on and overall time required to purify 96 samples compare to the manual purification method. Thus, the use of commercial bead beating tubes, RBB method and automatic purification system provided good quality fecal DNA with significantly higher throughput.

Disruption of bacterial cells to release DNA is completely dependent upon the type of cell structure, especially the cell wall. For an effective breakdown of cell wall, freezing and thawing prior to bead beating have also been very effective and results in higher yield of DNA along with the strong mechanical disruption (bead beating) method (Bahl, Bergström, and Licht 2012; Guo and Zhang 2013), but in this project no tests were conducted to test the effect of freezing as in clinical trials all samples need to be frozen anyway in order to process them in large batches. There are reports showing effects of freezing fecal samples on bacterial community, especially on two bacterial groups *Firmicutes* and *Bacteroidetes,* the ratio of which seems to increase amongst frozen samples when compared against non-frozen (Bahl, Bergström, and Licht 2012). Similar increase in ratio of in bacterial gene copies of *Firmicutes* and *Bacteroidetes* were observed in study performed by Metzler et al., while both DNA yield and microbial diversity were increased when extracted using bead beating methods (Metzler-Zebeli et al. 2016; Knudsen et al. 2016). Use of bead beating will primarily increase the DNA yield due to breakdown of microbial cell wall (Smith et al. 2011) from both gram positive and gram-negative bacteria. The RBB protocol implemented in this study involves two rounds of bead beating in lysis buffer, where the remaining intact cells from the first round undergo another round of bead beating after incubation at 95 °C, which altogether provides enough physical stress for a successful breakdown of most bacterial cells. Withdrawing the cell lysate after the first round of bead beating ensures no further damage to already released DNA. As previously shown for adult samples (Salonen 2010), RBB method facilitates efficient extraction of DNA from archaea, *Clostridium cluster* IV and other hard-to-lyse bacteria.

Subsequently we tested self-made wash buffers against commercial wash buffers used in purification of DNA with aim of circumventing the need to purchase extra bottles of expensive commercial wash buffers that are needed for each DNA purification run in 96 well format using the MagMAX Pathogen High Vol Duo protocol in KingFisher Flex automat. Determining the composition of wash buffers was another hurdle. Composition of commercial wash buffers were not disclosed by the company (MagMAX™) thus we went through the available literatures and finally one standard composition was devised and tested for its effectiveness. The composition included Sodium Acetate (Na-Ac), EthyleneDiamineTetraaceticAcid (EDTA), Tris Hydrochloride (Tris-HCl) and ethanol (EtOH). Including Sodium Acetate (Na-acetate) in wash buffer would help to separate higher molecular DNA and polysaccharides, neutralize base from denaturing step along with proteins, lipids and phosphate backbones, moreover it provides buffering over the use of normal salt like NaCl (Dellaporta, Wood, and Hicks 1983; Tan and Yiap 2009; OpenWetWare contributors 2016). The pH of Na-Ac was maintained at ~5.2 as suggested by different sources and to perform above spoken activities pH 5.2 was found to be the optimum one. Similarly, addition of EDTA in our buffer has already produced positive results and its use has fulfilled its purpose. Briefly, EDTA binds with Magnesium (Mg) ions (a co-factor for enzyme DNase) and prevent DNases to degrade DNA. Meanwhile use of Tris-HCl along with EDTA eventually helps to maintain the buffer's pH near to 8.0 restricting DNase's activity at pH 7.0 (Durst and Staples 1972; Dellaporta, Wood, and Hicks 1983; Yagi et al. 1996b; Somma and Querci 2004). Ethanol is commonly used for precipitating DNA by desalting (removal of carry over salts from previous steps) and minimizing solubility unlike when dissolved in water (Piškur and Rupprecht 1995; Zumbo 2012). Finally, three wash buffers were formulated, wash buffer I with and without EDTA and wash buffer II, pH of buffer was maintained between 7.5 and 8.0 like that of Magmax™ (commercial buffer).

Our tests revealed that the self-made wash buffers could replace expensive commercial wash buffers. DNA yields from commercial and self-made wash buffer were similar (Magmax- 10.24 µg/mg, HWI/II- 10.39 µg/mg and HWI_E/II- 10.48 µg/mg of feces). Additionally, based on qPCR-derived threshold cycle (Cq) values (Nechvatal et al. 2008), the PCR performance of the purified DNA preps were similar for Magmax™ and self-made wash buffer with EDTA (HWI_E and HWII) while DNA purified with self-made wash buffer without EDTA (HWI and HWII) gave inconstant results and failed to amplify with the highest template amount (250 ng). The only difference between the self-made wash buffers is presence of EDTA and as in the

presence of EDTA, both the DNA yield and PCR-compatibility of the preps were comparable to that of commercial wash buffer, we decided to use exclusively self-made wash buffers in the DNA extraction protocol.

For high throughput and accurate quantification of extracted DNA, fluorometry-based quantification was used, Qubit® (low scale) and Quant-iT™ (96 well format) over NanoDrop™. Basically, NanoDrop™ measures anything that can absorb 260 nm light like single and double stranded DNAs, RNAs, proteins and other constituents including contaminants, while both Qubit® and Quant-iT™ detects fluorescence from double stranded DNA only with sensitivity of detection down to 0.2 ng. Additionally, Quant-iT™ method qualifies for high throughput system as it measures DNA concentrations of 96 samples simultaneously within 30 minutes. Thus, the sensitive and accurate fluorometric methods Qubit® and Quant-iT™ were selected for quantification of fecal DNA as part of the novel sample processing protocol.

6.3. Amplification of abundant bacteria with qPCR and 16S rRNA gene amplicon sequencing from fecal DNA extracted using different bead beating tubes and purification methods

We evaluated the effectiveness of DNA extraction methods and their constituents using qPCR method to amplify abundant bacterial genera along with total bacteria (Salonen et al. 2010; De Gregoris et al. 2011). From children's fecal DNA, the combination of readymade tubes with RBB method, i.e. the one selected for the final protocol, gave ca. 0.3 log lower amplifications for total bacteria and all the tested bacterial groups (*Bacteroides, Clostridium and Bifidobacterium*) compared to readymade tubes with BB method, or self-made tubes with RBB method (Figure 11). Reasons behind such differences in qPCR amplifications remain unknown and due to limited time, further extensive testing were not performed. For adult's fecal DNA samples that were purified with manual system, there were no differences in amplifications amongst the tested extraction methods (BB and RBB) performed with readymade bead beating tubes.

Comparison of the qPCR results for *Bifidobacterium, Clostridium* and *Bacteroides* between manually purified adult samples and automatically purified infant's samples indicated that adult samples had lower bacterial copy numbers when equal amount of bacterial DNA was used in qPCR reaction. This is unexpected as adult samples are expected to have higher counts of *Clostridium* and *Bacteroides*. As the qPCR tests were designed to compare DNA extraction conditions within a sample type (adult or child), it remains unclear if the detected differences are biological or reflect the different PCR performance of manually and automatically extracted DNA. Manually purified adult's DNA using silica columns likely contains more PCR-inhibitors than the children's samples purified using automated system based on paramagnetic beads, possibly explaining the lower copy numbers (less efficient amplification) of adult samples.

To obtain a community-wide assessment of the different DNA extracts, 16S rRNA genes from children's fecal DNA (purified automatically) were amplified and sequenced perform qualitative evaluation for extracts resulting from the combination of readymade bead beating tubes and RBB method as part of the high-throughput DNA extraction and automatic purification system. The analyzed 16S rRNA V3-V4 amplicons revealed numerous genus level taxa that were grouped into the dominant phyla, i.e. 2 gram-positive phyla Actinobacteria and Firmicutes, and 2 gram-negative phyla Bacteroidetes and Proteobacteria. There were no differences in the abundancy pattern between the different extraction methods (with different tubes and methods), and the individual differences in the microbiota compositions were accurately reproduced with all the tested methods.

6.4. Fecal DNA extraction from adult and infant samples

The newly validated and optimized fecal DNA extraction system was then applied to process samples from a clinical trial. In total 634 fecal samples from mothers (n=342) and their infants (n=292) were extracted using new high-throughput extraction and purification system. The concentrations of the extracted DNA samples varied extensively depending on the individual and upon the amount of fecal samples available for the extraction. The samples collected using swabs gave the lowest amount of DNA (in average 0.058 ng/µl for mothers and -2.94 ng/µl for children. Although the average amount of fecal material used in extraction were very similar

(~300 mg) for both adults and children, the amount of extracted DNA was much less for the children; on average 0.5 µg per extraction from ≤ 3 month old infants and 3.7 µg per extraction from the adult samples. This likely reflects the ongoing colonization in infants (Moughan et al. 1992; Koenig et al. 2011; Matamoros et al. 2013; Munyaka, Khafipour, and Ghia 2014) and possibly also the different sample consistency.

Due to the lower DNA yield from children's fecal samples, RBB microbial DNA extraction protocol was modified after the pilot experiments to make children fecal samples more concentrated for the extraction procedure. To this end, all infant's fecal and swab samples were homogenized with 500 µl PBS instead of 1000 µl used in adult samples. In total half (~50.5%) of the extracted DNA had negative concentrations; ~10.2% of adult swabs and ~40.4% children samples including swabs. Before proceeding with the majority of the extractions, we performed 16S rRNA gene full length (~1500 bp) PCR for a subset of the extracted DNA samples with concentrations ranging from negatives to very high positive. The positive PCR results gave supportive evidence for the presence of enough DNA for molecular microbiota analysis in most samples and justified the use of the developed extraction and purification method. In defense of the negative concentrations, the use of 1 µl of the original DNA eluate and diluting it 200 times during quantification with Quant-iT™ might not have been enough to contain sufficient DNA molecules to produce detectable fluorescence from dsDNA and thus resulted in negative concentrations when compared against standards. The modification in the protocol by using 500 µl PBS for homogenization of infant's fecal samples, better DNA yields and concentrations were obtained. With the success after the protocol's modification it has been successfully implemented in HELMi study (large scale microbiota study) in Finland.

6.5. Prevalence and persistence *L. rhamnosus* GG among 3 month's old children and co-occurrence in mother-infant pairs

Lactobacilli are the most common probiotic bacteria, which possesses many health benefits (Lebeer, Vanderleyden, and De Keersmaecker 2008). Specifically *L. rhamnosus* GG (ATCC 53103) have become a widely studied probiotic bacterium and several studies (Doron, Snydman, and Gorbach 2005; Kankainen et al. 2009; Rasinkangas et al. 2014) have reported its persistence in human (children and adults) gut as well as health benefits. In this study we

studied its prevalence and persistence among 3 month's old children and its co-occurrence in mother-infant pairs sampled on the same days.

After completing the *L. rhamnosus* GG experiments on the selected samples using the sortase C specific end-point PCR (Table 11), it turned out that this assay is not fully specific for *L. rhamnosus* GG but instead amplifies DNA also from some other closely related lactobacilli (see section 6.7). After development of the new qPCR based assay, the 3 month samples from JORVI and FLORA studies were used to validate the new assay (Figure 21). Unfortunately, due to time limitation for this master's thesis project it was not feasible to re-run all the samples with the qPCR assay. Hence, the *L. rhamnosus* GG results are mainly discussed based on the results of the end-point PCR that was performed in all selected samples.

At the time of childbirth, ~39% of the mothers were positive for *L. rhamnosus* GG, while none of the infants were positive during the first two weeks of life. On the other hand, at 3 months, ~60% of the infants were positive while only ~14% of the mothers were positive. From this unsynchronized outcome it can be concluded that the mother is not the primary route for the neonate for acquiring *L. rhamnosus* GG. For this asynchrony other factors, such as the use of probiotic products and antibiotics as well as the low DNA yields from infant samples, come into play.

The intake of *L. rhamnosus* GG as a probiotic supplement is the most obvious source of this bacterium in the feces. During daily intake, adults secrete ca. 10^8 genome copies of *L. rhamnosus* GG per gram of feces (Kekkonen et al. 2008). In this project, although the use of probiotic products was recorded in the clinical trials from where a small subset of samples was used in this study, the data were not collected in such detail that it would enable accurate comparison of *L. rhamnosus* GG intake versus excretion, especially regarding the actual duration and timing of probiotic intake in relation to fecal sampling. Accordingly, the use of binary data (yes/no) on the probiotic use before fecal sampling (Table 2) proved to correspond poorly to the presence of *L. rhamnosus* GG in the fecal samples. For example, at 3 months, none of the mothers who took probiotics containing *L. rhamnosus* GG (5/28) were positive in the PCR. In the case of children, ~57% (16/28) were given probiotics containing *L. rhamnosus* GG and among them, ~62% (n=16) were positive.

With the reliable qPCR assay, more than every second (57%) 3-month-old infant tested positive for *L. rhamnosus* GG in the recent JORVI study. Collectively, these data reflect active probiotic use during perinatal period in the current society. On the other hand, as high as 30% of the 3-

month-old infants in the control group of the FLORA study, conducted over a decade ago, were also positive, although the control group is advised not to use any probiotic products immediately before and during the intervention period. Our results are in line with those of a probiotic intervention trial in adults; despite the 3-wk run-in with probiotic restriction, a detectable level of *L. rhamnosus* GG was detected in a subset of subjects at baseline (Kekkonen et al. 2008). This is because in Finland probiotic use is highly common (Saxelin 1997) and staple foods such as yoghurts can contain probiotic bacteria so that they are consumed unintentionally (Ouwehand, Saxelins, and Salminen 2004).

The fact that a single infant was not positive for *L. rhamnosus* GG before the week 3 after birth may have explanation related to microbial ecology. *L. rhamnosus* GG is a facultative anaerobe and a newborn's gut is inclined towards facultative aerobes before the pioneering aerobes consume the prevailing oxygen (Houghteling and Walker 2015). Hence, it likely takes a while for *L. rhamnosus* GG to flourish in a neonatal gut. In addition, Goldin et al. have reported that for a successful detection of *L. rhamnosus* GG it takes at least a week after the beginning of intervention in adults (Goldin et al. 1992; Guaraldi and Salvatori 2012). Meanwhile, from an intervention study conducted among 2 to 7 years old children, Korpela et al. have described the successful alteration in intestinal microbiota after uninterrupted consumption of *L. rhamnosus* GG for 7 months (Korpela, Salonen, Virta, Kumpu et al. 2016). Thus, it likely takes some time for *L. rhamnosus* GG to prevail in unfavorable infant's gut despite its reported ingestion at earlier time points.

Based on literature, transfer of *L. rhamnosus* GG from mother can also happen via mother's milk (Ballard and Morrow 2013) and in both studies (FLORA and JORVI) exclusively breastfed infants (~31%) were more often positive for *L. rhamnosus* GG compared to those who were formula-fed (~12%). Similarly, antibiotics have also been reported as influencing factor in colonization and prevalence pattern of microorganisms including on *L. rhamnosus* GG (Korpela, Salonen, Virta, Kekkonen et al. 2016; Korpela, Salonen, Virta, Kumpu et al. 2016; Langdon, Crook, and Dantas 2016). Following the antibiotics intervention in JORVI study, among 11 mothers from antibiotics group (n=14) who completed their antibiotics dosage before giving birth, 5 were positive for *L. rhamnosus* GG at birth while only 1 was positive at 3 months, and in non-antibiotics group (n=14), 5 mothers were positive at the time of birth and 3 at 3 months. In case of infants (n=28) who did not take any antibiotics until 3 months, none were positive for *L. rhamnosus* GG at birth while 17 became positive at the age of 3 months. With the limited information about only intravenous antibiotics used during parturition and

their use during different time points limits us to get to a conclusion. Only with detailed study of the effects of those antibiotics on *L. rhamnosus* GG could produce a valuable result.

Looking at the technical part, it should be noted that the comparison of the multiple timepoints from mother-infant pairs for *L. rhamnosus* GG carriage was based on the end-point PCR that turned out not to be fully specific from the targeted strain. Hence, while the data may contain false positives, we also cannot exclude that some infant samples from JORVI study remained below the detection level due to the low DNA yield, and hence are false negatives. In addition, studying prevalence of bacterium that is known to bind intestinal mucus (Kankainen et al. 2009; Rasinkangas et al. 2014) using fecal samples have also been questioned since Alander et al. have explained that *L. rhamnosus* GG persists for longer periods in the colonic mucosa layer even when not detected in feces (Alander et al. 1999). Overall, the focus area of this project was in the development of high throughput fecal sample processing system, its application for the first time for samples derived from a clinical trial, and confirmation that the resultant DNA is suitable for bacterial analysis. The results of *L. rhamnosus* GG analysis helped us to identify the points for further improvement, more specifically the need for higher input (which eventually gives more concentrated samples) of infant fecal samples for sufficient DNA yield, as well as the need for better method for detection of *L. rhamnosus* GG. Hence, a broader discussion and speculation on the factors that underlie the biological outcome of the *L. rhamnosus* GG results is out of the scope of this project and instead are justified for another study which has larger sampler size and that is free from technical confounders related e.g. to low DNA yield and assay specificity.

6.6. New specific identification and quantification real-time detection method to replace old unspecific method for identification of *L. rhamnosus* GG

A side project was also going on during this project, a new alternative approach was being developed for specific identification and quantification of *L. rhamnosus* GG that could replace old screening method (Table 11, sortase C specific end-point PCR). Incorrect identification of bacteria and outcome from any research is unreliable, Hamilton and Shah have also shown the methods that have misidentified different strains of bacteria including *L. rhamnosus* GG (Hamilton-Miller and Shah 2002). With known characteristics of *L. rhamnosus* GG's pili (containing mucus binding protein) and role of different sortases (SrtA and SrtC) (Kankainen

et al. 2009; Douillard et al. 2014) have led us to use SrtC gene in development of new detection and quantification method for *L. rhamnosus* GG. New specific real-time identification and quantification method (qPCR) for *L. rhamnosus* GG targets section containing partial genomic and sortase C gene, which proved to be a reliable and robust method instead of targeting only SrtC gene (Figure 20 and 21). Results confirmed from the blast hits of sequences proved this new method's specificity and reliability over old method. Old method's target gene produced multiple blast hits from *L. casei* and *L. paracasei* along with desired *L. rhamnosus* GG but new real-time detection method showed specificity to two strains of *L. rhamnosus* GG and a plasmid, which clearly is very specific to *L. rhamnosus* GG. Along with specific identification, quantification was also possible with this new method. This gives us enormous advantage over old PCR method. To understand the reliability of this new method same fecal DNA samples from 3 month's old children from JORVI (n=28) and FLORA (n=30) were analyzed and the outcome was shocking, which led us to question our results. Briefly, only 10 children were positive in case of FLORA samples while 18 were positive when analyzed with old method. In JORVI new method showed 16 positives than old method which gave 17 positives. With these differences, old unspecific and unreliable detection method was successfully replaced by new reliable method, and it has already been heavily used in new studies.

7. CONCLUSION

Our primary objective to establish an efficient fecal sample processing system for large-scale microbiome studies, including an up-scaled method for extraction of high quality microbial DNA using repeated bead beating method in high throughput manner (96 well format) have been achieved. Use of self-made lysis and washing buffers and automatic DNA purification method gave similar quality and quantity of microbial DNA to that of commercial one. This newly developed system has proven to yield high quality DNA from 96 samples within 6 to 8 hours with 3 to 4 hours of pre-processing, depending upon fecal sample types (liquid, semi-solid or solid). The protocol has been in use since its set up in de Vos/Salonen laboratory, and is being extensively used especially in a large cohort study known as HELMi (Health and Early Life Microbiome) in Finland. The use of self-made buffers during DNA extraction is not only economical but also ensures that the same protocol is available in longitudinal studies where the first and last samples can be sampled and processed several years apart, during which changes in the composition of commercial buffers can take place.

In this project, microbial DNA from 647 adults and children (≤ 3 months) fecal samples from JORVI study were successfully extracted with the newly established high throughput system. Of those, samples from birth to 3 months from 28 mother-infant pairs were randomly picked to test the suitability of the DNA extracts for PCR-based downstream analysis, and to set up a pilot study for understanding the prevalence of *L. rhamnosus* GG in neonates and their mothers. While we made some attempts to relate the results on *L. rhamnosus* GG to the available data on antibiotic and probiotic use, the *L. rhamnosus* GG results presented in this study must be interpreted with caution. This is due to the fact the data may contain both false positives and false negatives, because the applied end-point PCR method turned out to be somewhat unspecific, and some infant's samples yielded minute DNA yields, respectively. Hence, another study is required to properly study the prevalence of *L. rhamnosus* GG and factors affecting it, building on the methods developed and validated in this thesis, i.e. the specific qPCR assay as well as the DNA extraction method that was further optimized for infant samples by increasing the sample input. Hence, although the biological relevance of the current results on the prevalence of *L. rhamnosus* GG is limited, by identifying the pitfalls of the original study, set up the results of this project further contribute to the technical evaluation and validation of the targeted protocols.

8. References

Aagaard, K., J. Ma, K. M. Antony, R. Ganu, J. Petrosino, and J. Versalovic. 2014. "The Placenta Harbors a Unique Microbiome." Science Translational Medicine 6 (237): 237ra65.

Adlerberth, I. and AE Wold. 2009. "Establishment of the Gut Microbiota in Western Infants." *Acta Paediatrica* 98 (2): 229-238.

Agans, Richard, Laura Rigsbee, Harshavardhan Kenche, Sonia Michail, Harry J. Khamis, and Oleg Paliy. 2011. "Distal Gut Microbiota of Adolescent Children is Different from that of Adults." *FEMS Microbiology Ecology* 77 (2): 404-412.

Aguirre, Marisol and Koen Venema. 2015. "The use of Fecal Samples for Studying Human Obesity." *European Journal of Epidemiology* 30 (9): 1067-1069.

Akobeng, A. K., A. V. Ramanan, I. Buchan, and R. F. Heller. 2006. "Effect of Breast Feeding on Risk of Coeliac Disease: A Systematic Review and Meta-Analysis of Observational Studies." *Archives of Disease in Childhood* 91 (1): 39-43.

Akutsu, Jun-ichi, Yuriko Tojo, Osamu Segawa, Kimimichi Obata, Mina Okochi, Hideji Tajima, and Masafumi Yohda. 2004. "Development of an Integrated Automation System with a Magnetic Bead-mediated Nucleic Acid Purification Device for Genetic Analysis and Gene Manipulation." *Biotechnology and Bioengineering* 86 (6): 667-671.

Alander, M., R. Satokari, R. Korpela, M. Saxelin, T. Vilpponen-Salmela, T. Mattila-Sandholm, and A. von Wright. 1999. "Persistence of Colonization of Human Colonic Mucosa by a Probiotic Strain, Lactobacillus Rhamnosus GG, After Oral Consumption." *Applied and Environmental Microbiology* 65 (1): 351-354.

Algert, CS, A. McElduff, JM Morris, and CL Roberts. 2009. "Perinatal Risk Factors for Early Onset of Type 1 Diabetes in a 2000–2005 Birth Cohort." *Diabetic Medicine* 26 (12): 1193-1197.

Amplicon, PCR, PCR Clean-Up, and PCR Index. 2013.*16S Metagenomic Sequencing Library Preparation*.

Ardissone, Alexandria N., M. Diomel, Austin G. Davis-Richardson, Kevin T. Rechcigl, Nan Li, Jennifer C. Drew, Roberto Murgas-Torrazza, Renu Sharma, Mark L. Hudak, and Eric W. Triplett. 2014. "Meconium Microbiome Analysis Identifies Bacteria Correlated with Premature Birth." *PloS One* 9 (3): e90784.

Azad, MB, T. Konya, RR Persaud, DS Guttman, RS Chari, CJ Field, MR Sears, PJ Mandhane, SE Turvey, and P. Subbarao. 2016. "Impact of Maternal Intrapartum Antibiotics, Method of Birth and Breastfeeding on Gut Microbiota during the First Year of Life: A Prospective Cohort Study." *BJOG: An International Journal of Obstetrics & Gynaecology* 123 (6): 983-993.

Bäckhed, Fredrik, Josefine Roswall, Yangqing Peng, Qiang Feng, Huijue Jia, Petia Kovatcheva-Datchary, Yin Li, Yan Xia, Hailiang Xie, and Huanzi Zhong. 2015. "Dynamics and Stabilization of the Human Gut Microbiome during the First Year of Life." *Cell Host & Microbe* 17 (5): 690-703.

Bahl, Martin Iain, Anders Bergström, and Tine Rask Licht. 2012. "Freezing Fecal Samples Prior to DNA Extraction Affects the Firmicutes to Bacteroidetes Ratio Determined by Downstream Quantitative PCR Analysis." *FEMS Microbiology Letters* 329 (2): 193-197.

Ballard, O. and A. L. Morrow. 2013. "Human Milk Composition: Nutrients and Bioactive Factors." *Pediatric Clinics of North America* 60 (1): 49-74.

Barbosa, Márcio Ribeiro. 2013. "Chemical Composition and Formation of Human Feces-Problems and Solutions of Large Mergers Demographics in Developing Countries.10th International Symposium on Recent Advances in Environmental Health Research.

Bearfield, Caroline, Elizabeth S. Davenport, Vythil Sivapathasundaram, and Robert P. Allaker. 2002. "Possible Association between Amniotic Fluid Micro-organism Infection and Microflora in the Mouth." *BJOG: An International Journal of Obstetrics & Gynaecology* 109 (5): 527-533.

Beaumont, Michelle, Julia K. Goodrich, Matthew A. Jackson, Idil Yet, Emily R. Davenport, Sara Vieira-Silva, Justine Debelius, Tess Pallister, Massimo Mangino, and Jeroen Raes. 2016. "Heritable Components of the Human Fecal Microbiome are Associated with Visceral Fat." *Genome Biology* 17 (1): 189.

Behjati, S. and P. S. Tarpey. 2013. "What is Next Generation Sequencing?" *Archives of Disease in Childhood.Education and Practice Edition* 98 (6): 236-238.

Belkaid, Yasmine and Timothy W. Hand. 2014. "Role of the Microbiota in Immunity and Inflammation." *Cell* 157 (1): 121-141.

Benno, Y. 1985. "The Intestinal Microflora of Infants; Composition of Fecal Flora in Breast and Bottle Fed Infants." *The Pediatric Infectious Disease Journal* 4 (3): 308.

Bezirtzoglou, Eugenia, Arsenis Tsiotsias, and Gjalt W. Welling. 2011. "Microbiota Profile in Feces of Breast-and Formula-Fed Newborns by using Fluorescence in Situ Hybridization (FISH)." *Anaerobe* 17 (6): 478-482.

Biasucci, G., B. Benenati, L. Morelli, E. Bessi, and G. Boehm. 2008. "Cesarean Delivery may Affect the Early Biodiversity of Intestinal Bacteria." *The Journal of Nutrition* 138 (9): 1796S-1800S.

Blaut, M., MD Collins, GW Welling, J. Dore, J. Van Loo, and W. De Vos. 2002. "Molecular Biological Methods for Studying the Gut Microbiota: The EU Human Gut Flora Project." *British Journal of Nutrition* 87 (S2): S203-S211.

Boom, R., C. J. Sol, M. M. Salimans, C. L. Jansen, P. M. Wertheim-van Dillen, and J. van der Noordaa. 1990. "Rapid and Simple Method for Purification of Nucleic Acids." *Journal of Clinical Microbiology* 28 (3): 495-503.

Borre, Yuliya E., Rachel D. Moloney, Gerard Clarke, Timothy G. Dinan, and John F. Cryan. 2014. "The Impact of Microbiota on Brain and Behavior: Mechanisms & Therapeutic Potential." In *Microbial Endocrinology: The Microbiota-Gut-Brain Axis in Health and Disease*, 373-403.

Bron, Peter A., Peter Van Baarlen, and Michiel Kleerebezem. 2012. "Emerging Molecular Insights into the Interaction between Probiotics and the Host Intestinal Mucosa." *Nature Reviews.Microbiology* 10 (1): 66.

Call, Emma and Todd Klaenhammer. 2013. "Relevance and Application of Sortase and Sortase-Dependent Proteins in Lactic Acid Bacteria." *Frontiers in Microbiology* 4: 73.

Cattaneo, Adriano, Agneta Yngve, Berthold Koletzko, and Luis Ruiz Guzman. 2005. "Protection, Promotion and Support of Breast-Feeding in Europe: Current Situation." *Public Health Nutrition* 8 (1): 39-46.

Choo, Jocelyn M., Lex EX Leong, and Geraint B. Rogers. 2015. "Sample Storage Conditions significantly Influence Fecal Microbiome Profiles." *Scientific Reports* 5: 16350.

Claassen, Shantelle, Elloise du Toit, Mamadou Kaba, Clinton Moodley, Heather J. Zar, and Mark P. Nicol. 2013. "A Comparison of the Efficiency of Five Different Commercial DNA Extraction Kits for Extraction of DNA from Fecal Samples." *Journal of Microbiological Methods* 94 (2): 103-110.

Costea, Paul I., Georg Zeller, Shinichi Sunagawa, Eric Pelletier, Adriana Alberti, Florence Levenez, Melanie Tramontano, Marja Driessen, Rajna Hercog, and Ferris-Elias Jung. 2017. "Towards Standards for Human Fecal Sample Processing in Metagenomic Studies." *Nature Biotechnology* 35 (11): 1069.

Cox, Laura M., Shingo Yamanishi, Jiho Sohn, Alexander V. Alekseyenko, Jacqueline M. Leung, Ilseung Cho, Sungheon G. Kim, Huilin Li, Zhan Gao, and Douglas Mahana. 2014. "Altering the Intestinal Microbiota during a Critical Developmental Window has Lasting Metabolic Consequences." *Cell* 158 (4): 705-721.

Daliri, Eric Banan-Mwine, Shuai Wei, Deog H. Oh, and Byong H. Lee. 2017. "The Human Microbiome and Metabolomics: Current Concepts and Applications." *Critical Reviews in Food Science and Nutrition* 57 (16): 3565-3576.

De Filippo, C., D. Cavalieri, M. Di Paola, M. Ramazzotti, J. B. Poullet, S. Massart, S. Collini, G. Pieraccini, and P. Lionetti. 2010. "Impact of Diet in Shaping Gut Microbiota Revealed by a Comparative Study in Children from Europe and Rural Africa." *Proceedings of the National Academy of Sciences of the United States of America* 107 (33): 14691-14696.

De Gregoris, Tristano Bacchetti, Nick Aldred, Anthony S. Clare, and J. Grant Burgess. 2011. "Improvement of Phylum-and Class-Specific Primers for Real-Time PCR Quantification of Bacterial Taxa." *Journal of Microbiological Methods* 86 (3): 351-356.

de Tejada, Begoña Martinez. 2014. "Antibiotic use and Misuse during Pregnancy and Delivery: Benefits and Risks." *International Journal of Environmental Research and Public Health* 11 (8): 7993-8009.

Decker, E., G. Engelmann, A. Findeisen, P. Gerner, M. Laass, D. Ney, C. Posovszky, L. Hoy, and M. W. Hornef. 2010. "Cesarean Delivery is Associated with Celiac Disease but Not Inflammatory Bowel Disease in Children." *Pediatrics* 125 (6): e1433-40.

Dellaporta, Stephen L., Jonathan Wood, and James B. Hicks. 1983. "A Plant DNA Minipreparation: Version II." *Plant Molecular Biology Reporter* 1 (4): 19-21.

Doron, S., D. R. Snydman, and S. L. Gorbach. 2005. "Lactobacillus GG: Bacteriology and Clinical Applications." *Gastroenterology Clinics of North America* 34 (3): 483-98, ix.

Douillard, F. P., P. Rasinkangas, I. von Ossowski, J. Reunanen, A. Palva, and W. M. de Vos. 2014. "Functional Identification of Conserved Residues Involved in Lactobacillus Rhamnosus Strain GG Sortase Specificity and Pilus Biogenesis." *The Journal of Biological Chemistry* 289 (22): 15764-15775.

Durst, R. A. and B. R. Staples. 1972. "Tris-Tris-HCl: A Standard Buffer for use in the Physiologic pH Range." *Clinical Chemistry* 18 (3): 206-208.

Eckburg, P. B., E. M. Bik, C. N. Bernstein, E. Purdom, L. Dethlefsen, M. Sargent, S. R. Gill, K. E. Nelson, and D. A. Relman. 2005. "Diversity of the Human Intestinal Microbial Flora." *Science (New York, N.Y.)* 308 (5728): 1635-1638.

Fallani, M., D. Young, J. Scott, E. Norin, S. Amarri, R. Adam, M. Aguilera, et al. 2010. "Intestinal Microbiota of 6-Week-Old Infants Across Europe: Geographic Influence Beyond Delivery Mode, Breast-Feeding, and Antibiotics." *Journal of Pediatric Gastroenterology and Nutrition* 51 (1): 77-84.

Falony, G., M. Joossens, S. Vieira-Silva, J. Wang, Y. Darzi, K. Faust, A. Kurilshikov, et al. 2016. "Population-Level Analysis of Gut Microbiome Variation." *Science (New York, N.Y.)* 352 (6285): 560-564.

Fisher, Roberta M., Lee M. Henry, Charlie K. Cornwallis, E. Toby Kiers, and Stuart A. West. 2017. "The Evolution of Host-Symbiont Dependence." *Nature Communications* 8: 15973.

Franzreb, Matthias, Martin Siemann-Herzberg, Timothy J. Hobley, and Owen RT Thomas. 2006. "Protein Purification using Magnetic Adsorbent Particles." *Applied Microbiology and Biotechnology* 70 (5): 505-516.

Fujimoto, Shuji, Yoshiko Nakagami, and Fumiko Kojima. 2004. "Optimal Bacterial DNA Isolation Method using Bead-Beating Technique." *Memoirs Kyushu Univ Dep of Health Scis of Medical Sch* 3: 33-38.

Fung, Thomas C., Christine A. Olson, and Elaine Y. Hsiao. 2017. "Interactions between the Microbiota, Immune and Nervous Systems in Health and Disease." *Nature Neuroscience* 20 (2): 145.

Funkhouser, Lisa J. and Seth R. Bordenstein. 2013. "Mom Knows Best: The Universality of Maternal Microbial Transmission." *PLoS Biology* 11 (8): e1001631.

Goldin, Barry R., Sherwood L. Gorbach, Maija Saxelin, Susan Barakat, Lisa Gualtieri, and Seppo Salminen. 1992. "Survival ofLactobacillus Species (Strain GG) in Human Gastrointestinal Tract." *Digestive Diseases and Sciences* 37 (1): 121-128.

Gorbach, Sherwood L. 2000. "Probiotics and Gastrointestinal Health." *The American Journal of Gastroenterology* 95 (1): S2-S4.

Govender, Mershen, Yahya E. Choonara, Pradeep Kumar, Lisa C. du Toit, Sandy van Vuuren, and Viness Pillay. 2014. "A Review of the Advancements in Probiotic Delivery: Conventional Vs. Non-Conventional Formulations for Intestinal Flora Supplementation." *AAPS PharmSciTech* 15 (1): 29-43.

Guandalini, Stefano, Licia Pensabene, Mona Abu Zikri, Jorge Amil Dias, Luigi Gobio Casali, Hans Hoekstra, Sanja Kolacek, Karin Massar, Dusanka Micetic–Turk, and Alexandra Papadopoulou. 2000. "Lactobacillus GG Administered in Oral Rehydration Solution to Children with Acute Diarrhea: A Multicenter European Trial." *Journal of Pediatric Gastroenterology and Nutrition* 30 (1): 54-60.

Guaraldi, F. and G. Salvatori. 2012. "Effect of Breast and Formula Feeding on Gut Microbiota Shaping in Newborns." *Frontiers in Cellular and Infection Microbiology* 2: 94.

Guarner, Francisco and Juan-R Malagelada. 2003. "Gut Flora in Health and Disease." *The Lancet* 361 (9356): 512-519.

Guo, Feng and Tong Zhang. 2013. "Biases during DNA Extraction of Activated Sludge Samples Revealed by High Throughput Sequencing." *Applied Microbiology and Biotechnology* 97 (10): 4607-4616.

Gurson, Natalie. "**when do I use Sanger Sequencing Vs NGS?**" ThermoFisher Scientific, accessed 04/12, 2018, https://www.thermofisher.com/blog/behindthebench/when-do-i-use-sanger-sequencing-vs-ngs-seq-it-out-7/.

Hamilton-Miller, JMT and S. Shah. 2002. "Deficiencies in Microbiological Quality and Labelling of Probiotic Supplements." *International Journal of Food Microbiology* 72 (1): 175-176.

Hart, Marcia L., Alexandra Meyer, Philip J. Johnson, and Aaron C. Ericsson. 2015. "Comparative Evaluation of DNA Extraction Methods from Feces of Multiple Host Species for Downstream Next-Generation Sequencing." *PloS One* 10 (11): e0143334.

Hatakka, K., E. Savilahti, A. Ponka, J. H. Meurman, T. Poussa, L. Nase, M. Saxelin, and R. Korpela. 2001. "Effect of Long Term Consumption of Probiotic Milk on Infections in Children Attending Day Care Centres: Double Blind, Randomised Trial." *BMJ (Clinical Research Ed.)* 322 (7298): 1327.

Hayashi, Hidenori, Mitsuo Sakamoto, and Yoshimi Benno. 2002. "Phylogenetic Analysis of the Human Gut Microbiota using 16S rDNA Clone Libraries and Strictly Anaerobic Culture-based Methods." *Microbiology and Immunology* 46 (8): 535-548.

Heid, C. A., J. Stevens, K. J. Livak, and P. M. Williams. 1996. "Real Time Quantitative PCR." *Genome Research* 6 (10): 986-994.

Hemarajata, Peera and James Versalovic. 2013. "Effects of Probiotics on Gut Microbiota: Mechanisms of Intestinal Immunomodulation and Neuromodulation." *Therapeutic Advances in Gastroenterology* 6 (1): 39-51.

Hiergeist, Andreas, Joachim Gläsner, Udo Reischl, and André Gessner. 2015. "Analyses of Intestinal Microbiota: Culture Versus Sequencing." *ILAR Journal* 56 (2): 228-240.

Holland, P. M., R. D. Abramson, R. Watson, and D. H. Gelfand. 1991. "Detection of Specific Polymerase Chain Reaction Product by Utilizing the 5'----3' Exonuclease Activity of Thermus Aquaticus DNA Polymerase." *Proceedings of the National Academy of Sciences of the United States of America* 88 (16): 7276-7280.

Hotel, Amerian Córdoba Park and A. Cordoba. 2001. "Health and Nutritional Properties of Probiotics in Food Including Powder Milk with Live Lactic Acid Bacteria." *Prevention* 5 (1).

Houghteling, P. D. and W. A. Walker. 2015. "Why is Initial Bacterial Colonization of the Intestine Important to Infants' and Children's Health?" *Journal of Pediatric Gastroenterology and Nutrition* 60 (3): 294-307.

Hu, Jianzhong, Yoko Nomura, Ali Bashir, Heriberto Fernandez-Hernandez, Steven Itzkowitz, Zhiheng Pei, Joanne Stone, Holly Loudon, and Inga Peter. 2013. "Diversified Microbiota of Meconium is Affected by Maternal Diabetes Status." *PLoS One* 8 (11): e78257.

Huh, S. Y., S. L. Rifas-Shiman, C. A. Zera, J. W. Edwards, E. Oken, S. T. Weiss, and M. W. Gillman. 2012. "Delivery by Caesarean Section and Risk of Obesity in Preschool Age Children: A Prospective Cohort Study." *Archives of Disease in Childhood* 97 (7): 610-616.

Human Microbiome Project Consortium. 2012. "Structure, Function and Diversity of the Healthy Human Microbiome." *Nature* 486 (7402): 207-214.

Hviid, A., H. Svanstrom, and M. Frisch. 2011. "Antibiotic use and Inflammatory Bowel Diseases in Childhood." *Gut* 60 (1): 49-54.

Isolauri, Erika, Philip M. Sherman, and W. Allan Walker. 2017. *Intestinal Microbiome: Functional Aspects in Health and Disease: 88th Nestlé Nutrition Institute Workshop, Playa Del Carmen, September 2016* Karger Medical and Scientific Publishers.

Jandhyala, S. M., R. Talukdar, C. Subramanyam, H. Vuyyuru, M. Sasikala, and D. Nageshwar Reddy. 2015. "Role of the Normal Gut Microbiota." *World Journal of Gastroenterology* 21 (29): 8787-8803.

Jenmalm, Maria C. 2017. "The Mother–offspring Dyad: Microbial Transmission, Immune Interactions and Allergy Development." *Journal of Internal Medicine*.

Jiménez, Esther, Leonides Fernández, María L. Marín, Rocío Martín, Juan M. Odriozola, Carmen Nueno-Palop, Arjan Narbad, Mónica Olivares, Jordi Xaus, and Juan M. Rodríguez. 2005. "Isolation of Commensal Bacteria from Umbilical Cord Blood of Healthy Neonates Born by Cesarean Section." *Current Microbiology* 51 (4): 270-274.

Jones, G. F., G. E. Ward, M. P. Murtaugh, G. Lin, and C. J. Gebhart. 1993. "Enhanced Detection of Intracellular Organism of Swine Proliferative Enteritis, Ileal Symbiont Intracellularis, in Feces by Polymerase Chain Reaction." *Journal of Clinical Microbiology* 31 (10): 2611-2615.

Kankainen, M., L. Paulin, S. Tynkkynen, I. von Ossowski, J. Reunanen, P. Partanen, R. Satokari, et al. 2009. "Comparative Genomic Analysis of Lactobacillus Rhamnosus GG Reveals Pili Containing a Human- Mucus Binding Protein." *Proceedings of the National Academy of Sciences of the United States of America* 106 (40): 17193-17198.

Kekkonen, R. A., N. Lummela, H. Karjalainen, S. Latvala, S. Tynkkynen, S. Jarvenpaa, H. Kautiainen, I. Julkunen, H. Vapaatalo, and R. Korpela. 2008. "Probiotic Intervention has Strain-Specific Anti-Inflammatory Effects in Healthy Adults." *World Journal of Gastroenterology* 14 (13): 2029-2036.

Klindworth, Anna, Elmar Pruesse, Timmy Schweer, Jörg Peplies, Christian Quast, Matthias Horn, and Frank Oliver Glöckner. 2013. "Evaluation of General 16S Ribosomal RNA Gene PCR Primers for Classical and Next-Generation Sequencing-Based Diversity Studies." *Nucleic Acids Research* 41 (1): e1-e1.

Knudsen, Berith E., Lasse Bergmark, Patrick Munk, Oksana Lukjancenko, Anders Priemé, Frank M. Aarestrup, and Sünje J. Pamp. 2016. "Impact of Sample Type and DNA Isolation Procedure on Genomic Inference of Microbiome Composition." *Msystems* 1 (5): e00095-16.

Koenig, J. E., A. Spor, N. Scalfone, A. D. Fricker, J. Stombaugh, R. Knight, L. T. Angenent, and R. E. Ley. 2011. "Succession of Microbial Consortia in the Developing Infant Gut Microbiome." *Proceedings of the National Academy of Sciences of the United States of America* 108 Suppl 1: 4578-4585.

Kok, R. G., A. de Waal, F. Schut, G. W. Welling, G. Weenk, and K. J. Hellingwerf. 1996. "Specific Detection and Analysis of a Probiotic Bifidobacterium Strain in Infant Feces." *Applied and Environmental Microbiology* 62 (10): 3668-3672.

Korpela, Katri, Anne Salonen, Lauri J. Virta, Riina A. Kekkonen, Kristoffer Forslund, Peer Bork, and Willem M. De Vos. 2016. "Intestinal Microbiome is Related to Lifetime Antibiotic use in Finnish Pre-School Children." *Nature Communications* 7: ncomms10410.

Korpela, Katri, Anne Salonen, Lauri J. Virta, Minna Kumpu, Riina A. Kekkonen, and Willem M. De Vos. 2016. "Lactobacillus Rhamnosus GG Intake Modifies Preschool Children's Intestinal Microbiota, Alleviates Penicillin-Associated Changes, and Reduces Antibiotic Use." *PloS One* 11 (4): e0154012.

Kozyrskyj, Anita L., Pierre Ernst, and Allan B. Becker. 2007. "Increased Risk of Childhood Asthma from Antibiotic use in Early Life." *CHEST Journal* 131 (6): 1753-1759.

Kramer, Michael S. and Ritsuko Kakuma. 2004. "The Optimal Duration of Exclusive Breastfeeding." In *Protecting Infants through Human Milk*, 63-77.

Kukkonen, Kaarina, Erkki Savilahti, Tari Haahtela, Kaisu Juntunen-Backman, Riitta Korpela, Tuija Poussa, Tuula Tuure, and Mikael Kuitunen. 2007. "Probiotics and Prebiotic Galacto-Oligosaccharides in the Prevention of Allergic Diseases: A Randomized, Double-Blind, Placebo-Controlled Trial." *Journal of Allergy and Clinical Immunology* 119 (1): 192-198.

Langdon, Amy, Nathan Crook, and Gautam Dantas. 2016. "The Effects of Antibiotics on the Microbiome Throughout Development and Alternative Approaches for Therapeutic Modulation." *Genome Medicine* 8 (1): 39.

Lauber, Christian L., Nicholas Zhou, Jeffrey I. Gordon, Rob Knight, and Noah Fierer. 2010. "Effect of Storage Conditions on the Assessment of Bacterial Community Structure in Soil and Human-Associated Samples." *FEMS Microbiology Letters* 307 (1): 80-86.

Lay, C., L. Rigottier-Gois, K. Holmstrom, M. Rajilic, E. E. Vaughan, W. M. de Vos, M. D. Collins, et al. 2005. "Colonic Microbiota Signatures Across Five Northern European Countries." *Applied and Environmental Microbiology* 71 (7): 4153-4155.

Lebeer, S., J. Vanderleyden, and S. C. De Keersmaecker. 2008. "Genes and Molecules of Lactobacilli Supporting Probiotic Action." *Microbiology and Molecular Biology Reviews : MMBR* 72 (4): 728-64, Table of Contents.

Lederberg, Joshua and Alexa T. McCray. 2001. "Ome SweetOmics--A Genealogical Treasury of Words." *The Scientist* 15 (7): 8-8.

Ledger, William J. and Martin J. Blaser. 2013. "Are we using Too Many Antibiotics during Pregnancy?" *BJOG: An International Journal of Obstetrics & Gynaecology* 120 (12): 1450-1452.

Lee, Linda G., Charles R. Connell, and Will Bloch. 1993. "Allelic Discrimination by Nick-Translation PCR with Fluorgenic Probes." *Nucleic Acids Research* 21 (16): 3761-3766.

Lesley, Scott A. 2001. "High-Throughput Proteomics: Protein Expression and Purification in the Postgenomic World." *Protein Expression and Purification* 22 (2): 159-164.

Lyons, Philip P., James F. Turnbull, Karl A. Dawson, and Margaret Crumlish. 2017. "Exploring the Microbial Diversity of the Distal Intestinal Lumen and Mucosa of Farmed Rainbow Trout Oncorhynchus Mykiss (Walbaum) using Next Generation Sequencing (NGS)." *Aquaculture Research* 48 (1): 77-91.

Mackie, R. I., A. Sghir, and H. R. Gaskins. 1999. "Developmental Microbial Ecology of the Neonatal Gastrointestinal Tract." *The American Journal of Clinical Nutrition* 69 (5): 1035S-1045S.

Macpherson, Andrew J., Emma Slack, Markus B. Geuking, and Kathy D. McCoy. 2009. "The Mucosal Firewalls Against Commensal Intestinal Microbes¨, Seminars in immunopathology, Springer, 31(2): 145-149.

Mahony, Jennifer, Gabriele A. Lugli, Douwe van Sinderen, and Marco Ventura. 2018. "Impact of Gut-Associated Bifidobacteria and their Phages on Health: Two Sides of the Same Coin?" *Applied Microbiology and Biotechnology*: 1-9.

Matamoros, Sebastien, Christele Gras-Leguen, Françoise Le Vacon, Gilles Potel, and Marie-France de La Cochetiere. 2013. "Development of Intestinal Microbiota in Infants and its Impact on Health." *Trends in Microbiology* 21 (4): 167-173.

Matsuki, T., K. Watanabe, J. Fujimoto, Y. Miyamoto, T. Takada, K. Matsumoto, H. Oyaizu, and R. Tanaka. 2002. "Development of 16S rRNA-Gene-Targeted Group-Specific Primers for the Detection and Identification of Predominant Bacteria in Human Feces." *Applied and Environmental Microbiology* 68 (11): 5445-5451.

McOrist, Alexandra L., Michelle Jackson, and Anthony R. Bird. 2002. "A Comparison of Five Methods for Extraction of Bacterial DNA from Human Fecal Samples." *Journal of Microbiological Methods* 50 (2): 131-139.

Metchnikoff, Elie and Sir Peter Chalmers MITCHELL. 1907. *The Prolongation of Life: Optimistic Studies... the English Translation Edited by P. Chalmers Mitchell* London.

Metzler-Zebeli, Barbara U., Peadar G. Lawlor, Elizabeth Magowan, and Qendrim Zebeli. 2016. "Effect of Freezing Conditions on Fecal Bacterial Composition in Pigs." *Animals* 6 (3): 18.

Mirsepasi, Hengameh, Søren Persson, Carsten Struve, Lee OB Andersen, Andreas M. Petersen, and Karen A. Krogfelt. 2014. "Microbial Diversity in Fecal Samples Depends on DNA Extraction Method: easyMag DNA Extraction Compared to QIAamp DNA Stool Mini Kit Extraction." *BMC Research Notes* 7 (1): 50.

Moles, Laura, Marta Gómez, Hans Heilig, Gerardo Bustos, Susana Fuentes, Willem de Vos, Leónides Fernández, Juan M. Rodríguez, and Esther Jiménez. 2013. "Bacterial Diversity in Meconium of Preterm Neonates and Evolution of their Fecal Microbiota during the First Month of Life." *PloS One* 8 (6): e66986.

Moughan, PJ, MJ Birtles, PD Cranwell, WC Smith, and M. Pedraza. 1992. "The Piglet as a Model Animal for Studying Aspects of Digestion and Absorption in Milk-Fed Human Infants." In *Nutritional Triggers for Health and in Disease*. Vol. 67, 40-113.

Mueller, Noel T., Elizabeth Bakacs, Joan Combellick, Zoya Grigoryan, and Maria G. Dominguez-Bello. 2015. "The Infant Microbiome Development: Mom Matters." *Trends in Molecular Medicine* 21 (2): 109-117.

Mueller, Noel T., Robin Whyatt, Lori Hoepner, Sharon Oberfield, M. Gloria Dominguez-Bello, EM Widen, A. Hassoun, F. Perera, and A. Rundle. 2015. "Prenatal Exposure to Antibiotics, Cesarean Section and Risk of Childhood Obesity." *International Journal of Obesity* 39 (4): 665-670.

Munyaka, P. M., E. Khafipour, and J. E. Ghia. 2014. "External Influence of Early Childhood Establishment of Gut Microbiota and Subsequent Health Implications." *Frontiers in Pediatrics* 2: 109.

Nadkarni, Mangala A., F. Elizabeth Martin, Nicholas A. Jacques, and Neil Hunter. 2002. "Determination of Bacterial Load by Real-Time PCR using a Broad-Range (Universal) Probe and Primers Set." *Microbiology* 148 (1): 257-266.

Näse, Leena, K. Hatakka, Erkki Savilahti, M. Saxelin, A. Pönkä, Tuija Poussa, R. Korpela, and Jukka H. Meurman. 2001. "Effect of Long–term Consumption of a Probiotic Bacterium, Lactobacillus Rhamnosus GG, in Milk on Dental Caries and Caries Risk in Children." *Caries Research* 35 (6): 412-420.

Nechvatal, Jordan M., Jeffrey L. Ram, Marc D. Basson, Phanramphoei Namprachan, Stephanie R. Niec, Kawsar Z. Badsha, Larry H. Matherly, Adhip PN Majumdar, and Ikuko Kato. 2008. "Fecal Collection, Ambient Preservation, and DNA Extraction for PCR Amplification of Bacterial and Human Markers from Human Feces." *Journal of Microbiological Methods* 72 (2): 124-132.

NIH HMP Working Group, J. Peterson, S. Garges, M. Giovanni, P. McInnes, L. Wang, J. A. Schloss, et al. 2009. "The NIH Human Microbiome Project." *Genome Research* 19 (12): 2317-2323.

Oftedal, Olav T. 2002. "The Mammary Gland and its Origin during Synapsid Evolution." *Journal of Mammary Gland Biology and Neoplasia* 7 (3): 225-252.

Ogra, P. L. and Sr Welliver RC. 2008. "Effects of Early Environment on Mucosal Immunologic Homeostasis, Subsequent Immune Responses and Disease Outcome." *Nestle Nutrition Workshop Series.Paediatric Programme* 61: 145-181.

Olsen, Gary J., David J. Lane, Stephen J. Giovannoni, Norman R. Pace, and David A. Stahl. 1986. "Microbial Ecology and Evolution: A Ribosomal RNA Approach." *Annual Reviews in Microbiology* 40 (1): 337-365.

OpenWetWare contributors. "Qiagen Buffers." OpenWetWare, accessed September 18, 2017, http://www.openwetware.org/index.php?title=Qiagen_Buffers&oldid=943439.

Othman, Mohammad, JP Neilson, and Z. Alfirevic. 2007. "Probiotics for Preventing Preterm Labour." *Cochrane Database Syst.Rev* 1.

Ouwehand, Arthur C., Maija Saxelins, and Seppo Salminen. 2004. "Phenotypic Differences between Commercial Lactobacillus Rhamnosus GG and L. Rhamnosus Strains Recovered from Blood." *Clinical Infectious Diseases* 39 (12): 1858-1860.

Pareek, Chandra Shekhar, Rafal Smoczynski, and Andrzej Tretyn. 2011. "Sequencing Technologies and Genome Sequencing." *Journal of Applied Genetics* 52 (4): 413-435.

Parvez, S., Karim A. Malik, S. Ah Kang, and H-Y Kim. 2006. "Probiotics and their Fermented Food Products are Beneficial for Health." *Journal of Applied Microbiology* 100 (6): 1171-1185.

Petherick, AnnA. 2010. "Development: Mother's Milk: A Rich Opportunity." *Nature* 468 (7327): S5-S7.

Pflughoeft, Kathryn J. and James Versalovic. 2012. "Human Microbiome in Health and Disease." *Annual Review of Pathology: Mechanisms of Disease* 7: 99-122.

Picard, C., C. Ponsonnet, E. Paget, X. Nesme, and P. Simonet. 1992. "Detection and Enumeration of Bacteria in Soil by Direct DNA Extraction and Polymerase Chain Reaction." *Applied and Environmental Microbiology* 58 (9): 2717-2722.

Piškur, Jure and Allan Rupprecht. 1995. "Aggregated DNA in Ethanol Solution." *FEBS Letters* 375 (3): 174-178.

Qin, J., R. Li, J. Raes, M. Arumugam, K. S. Burgdorf, C. Manichanh, T. Nielsen, et al. 2010. "A Human Gut Microbial Gene Catalogue Established by Metagenomic Sequencing." *Nature* 464 (7285): 59-65.

Rajilić-Stojanović, Mirjana, Hans GHJ Heilig, Sebastian Tims, Erwin G. Zoetendal, and Willem M. Vos. 2013. "Long-term Monitoring of the Human Intestinal Microbiota Composition." *Environmental Microbiology* 15 (4): 1146-1159.

Rajilić-Stojanović, Mirjana and Willem M. de Vos. 2014. "The First 1000 Cultured Species of the Human Gastrointestinal Microbiota." *FEMS Microbiology Reviews* 38 (5): 996-1047.

Rasinkangas, P., J. Reunanen, F. P. Douillard, J. Ritari, V. Uotinen, A. Palva, and W. M. de Vos. 2014. "Genomic Characterization of Non-Mucus-Adherent Derivatives of Lactobacillus Rhamnosus GG Reveals Genes Affecting Pilus Biogenesis." *Applied and Environmental Microbiology* 80 (22): 7001-7009.

Rautava, S. 2017. "Microbial Composition of the Initial Colonization of Newborns." *Nestle Nutrition Institute Workshop Series* 88: 11-21.

Rautava, S., M. C. Collado, S. Salminen, and E. Isolauri. 2012. "Probiotics Modulate Host-Microbe Interaction in the Placenta and Fetal Gut: A Randomized, Double-Blind, Placebo-Controlled Trial." *Neonatology* 102 (3): 178-184.

Reunanen, J., I. von Ossowski, A. P. Hendrickx, A. Palva, and W. M. de Vos. 2012. "Characterization of the SpaCBA Pilus Fibers in the Probiotic Lactobacillus Rhamnosus GG." *Applied and Environmental Microbiology* 78 (7): 2337-2344.

Reuter, Jason A., Damek V. Spacek, and Michael P. Snyder. 2015. "High-Throughput Sequencing Technologies." *Molecular Cell* 58 (4): 586-597.

Rinttilä, T., A. Kassinen, E. Malinen, L. Krogius, and A. Palva. 2004. "Development of an Extensive Set of 16S rDNA-targeted Primers for Quantification of Pathogenic and Indigenous Bacteria in Fecal Samples by Real-time PCR." *Journal of Applied Microbiology* 97 (6): 1166-1177.

Rodríguez, Juan Miguel, Kiera Murphy, Catherine Stanton, R. Paul Ross, Olivia I. Kober, Nathalie Juge, Ekaterina Avershina, Knut Rudi, Arjan Narbad, and Maria C. Jenmalm. 2015. "The Composition of the Gut Microbiota Throughout Life, with an Emphasis on Early Life." *Microbial Ecology in Health and Disease* 26 (1): 26050.

Rothschild, Daphna, Omer Weissbrod, Elad Barkan, Alexander Kurilshikov, Tal Korem, David Zeevi, Paul I. Costea, et al. 2018. "Environment Dominates Over Host Genetics in Shaping Human Gut Microbiota." . http://dx.doi.org/10.1038/nature25973.

Rowland, Ian, Glenn Gibson, Almut Heinken, Karen Scott, Jonathan Swann, Ines Thiele, and Kieran Tuohy. 2017. "Gut Microbiota Functions: Metabolism of Nutrients and Other Food Components." *European Journal of Nutrition*: 1-24.

Šafařík, Ivo, Lucie Ptáčková, and Mirka Šafaříková. 2001. "Large-Scale Separation of Magnetic Bioaffinity Adsorbents." *Biotechnology Letters* 23 (23): 1953-1956.

Salonen, Anne, Janne Nikkilä, Jonna Jalanka-Tuovinen, Outi Immonen, Mirjana Rajilić-Stojanović, Riina A. Kekkonen, Airi Palva, and Willem M. de Vos. 2010. "Comparative Analysis of Fecal DNA Extraction Methods with Phylogenetic Microarray: Effective Recovery of Bacterial and Archaeal DNA using Mechanical Cell Lysis." *Journal of Microbiological Methods* 81 (2): 127-134.

Sanger, F., S. Nicklen, and A. R. Coulson. 1977. "DNA Sequencing with Chain-Terminating Inhibitors." *Proceedings of the National Academy of Sciences of the United States of America* 74 (12): 5463-5467.

Santiago, Alba, Suchita Panda, Griet Mengels, Xavier Martinez, Fernando Azpiroz, Joel Dore, Francisco Guarner, and Chaysavanh Manichanh. 2014. "Processing Fecal Samples: A Step Forward for Standards in Microbial Community Analysis." *BMC Microbiology* 14 (1): 112.

Saxelin, Maija. 1997. "Lactobacillus GG—a Human Probiotic Strain with Thorough Clinical Documentation." *Food Reviews International* 13 (2): 293-313.

Schuster, Stephan C. 2008. "Next-Generation Sequencing Transforms Today's Biology." *Nature Methods* 5 (1): 16.

Section on Breastfeeding. 2012. "Breastfeeding and the use of Human Milk." *Pediatrics* 129 (3): e827-41.

Sekirov, Inna, Shannon L. Russell, L. Caetano M. Antunes, and B. Brett Finlay. 2010. "Gut Microbiota in Health and Disease." *Physiological Reviews* 90 (3): 859-904.

Sender, Ron, Shai Fuchs, and Ron Milo. 2016. "Revised Estimates for the Number of Human and Bacteria Cells in the Body." *PLoS Biology* 14 (8): e1002533.

Shaw, Alexander G., Kathleen Sim, Elizabeth Powell, Emma Cornwell, Teresa Cramer, Zoë E. McClure, Ming-Shi Li, and J. Simon Kroll. 2016. "Latitude in Sample Handling and Storage for Infant Fecal Microbiota Studies: The Elephant in the Room?" *Microbiome* 4 (1): 40.

Smith, B., N. Li, A. S. Andersen, H. C. Slotved, and K. A. Krogfelt. 2011. "Optimising Bacterial DNA Extraction from Fecal Samples: Comparison of Three Methods." *The Open Microbiology Journal* 5: 14-17.

Somma, M. and M. Querci. 2004. "The Analysis of Food Samples for the Presence of Genetically Modified Organisms. Session 4. Extraction and Purification of DNA." *European Commission Joint Research Centre*.

Stark, Pamela L. and A. Lee. 1982. "The Microbial Ecology of the Large Bowel of Breastfed and Formula-Fed Infants during the First Year of Life." *Journal of Medical Microbiology* 15 (2): 189-203.

Stokholm, J., S. Schjørring, CE Eskildsen, L. Pedersen, AL Bischoff, N. Følsgaard, CG Carson, BLK Chawes, K. Bønnelykke, and A. Mølgaard. 2014. "Antibiotic use during Pregnancy Alters the Commensal Vaginal Microbiota." *Clinical Microbiology and Infection* 20 (7): 629-635.

Suau, A., R. Bonnet, M. Sutren, J. J. Godon, G. R. Gibson, M. D. Collins, and J. Dore. 1999. "Direct Analysis of Genes Encoding 16S rRNA from Complex Communities Reveals Many Novel Molecular Species within the Human Gut." *Applied and Environmental Microbiology* 65 (11): 4799-4807.

Tan, S. C. and B. C. Yiap. 2009. "DNA, RNA, and Protein Extraction: The Past and the Present." *Journal of Biomedicine & Biotechnology* 2009: 574398.

Tanaka, M. and J. Nakayama. 2017. "Development of the Gut Microbiota in Infancy and its Impact on Health in Later Life." *Allergology International: Official Journal of the Japanese Society of Allergology* 66 (4): 515-522.

Tanca, Alessandro, Marcello Abbondio, Antonio Palomba, Cristina Fraumene, Valeria Manghina, Francesco Cucca, Edoardo Fiorillo, and Sergio Uzzau. 2017. "Potential and Active Functions in the Gut Microbiota of a Healthy Human Cohort." *Microbiome* 5 (1): 79.

Tappenden, Kelly A. and Andrew S. Deutsch. 2007. "The Physiological Relevance of the Intestinal Microbiota-Contributions to Human Health." *Journal of the American College of Nutrition* 26 (6): 679S-683S.

Tedjo, Danyta I., Daisy MAE Jonkers, Paul H. Savelkoul, Ad A. Masclee, Niels van Best, Marieke J. Pierik, and John Penders. 2015. "The Effect of Sampling and Storage on the Fecal Microbiota Composition in Healthy and Diseased Subjects." *PloS One* 10 (5): e0126685.

Thomas, Carissa M. and James Versalovic. 2010. "Probiotics-Host Communication: Modulation of Signaling Pathways in the Intestine." *Gut Microbes* 1 (3): 148-163.

Trasande, L., J. Blustein, M. Liu, E. Corwin, L. M. Cox, and M. J. Blaser. 2013. "Infant Antibiotic Exposures and Early-Life Body Mass." *International Journal of Obesity (2005)* 37 (1): 16-23.

Turnbaugh, P. J., R. E. Ley, M. Hamady, C. M. Fraser-Liggett, R. Knight, and J. I. Gordon. 2007. "The Human Microbiome Project." *Nature* 449 (7164): 804-810.

Ursell, Luke K., Jessica L. Metcalf, Laura Wegener Parfrey, and Rob Knight. 2012. "Defining the Human Microbiome." *Nutrition Reviews* 70 (suppl_1): S38-S44.

Van der Waaij, D. 1983. "Colonization Pattern of the Digestive Tract by Potentially Pathogenic Microorganisms: Colonization-Controlling Mechanisms and Consequences for Antibiotic Treatment." *Infection* 11 (2): S90-S92.

Van Dijk, Erwin L., Hélène Auger, Yan Jaszczyszyn, and Claude Thermes. 2014. "Ten Years of Next-Generation Sequencing Technology." *Trends in Genetics* 30 (9): 418-426.

Vaughan, Elaine E., Hans GHJ Heilig, Erwin G. Zoetendal, Reetta Satokari, J. Kevin Collins, Antoon DL Akkermans, and Willem M. de Vos. 1999. "Molecular Approaches to Study Probiotic Bacteria." *Trends in Food Science & Technology* 10 (12): 400-404.

Vélez, Mónica Perea, Sigrid CJ De Keersmaecker, and Jos Vanderleyden. 2007. "Adherence Factors of Lactobacillus in the Human Gastrointestinal Tract." *FEMS Microbiology Letters* 276 (2): 140-148.

von Ossowski, I., J. Reunanen, R. Satokari, S. Vesterlund, M. Kankainen, H. Huhtinen, S. Tynkkynen, S. Salminen, W. M. de Vos, and A. Palva. 2010. "Mucosal Adhesion Properties of the Probiotic Lactobacillus Rhamnosus GG SpaCBA and SpaFED Pilin Subunits." *Applied and Environmental Microbiology* 76 (7): 2049-2057.

Walker, Alan W., Jennifer C. Martin, Paul Scott, Julian Parkhill, Harry J. Flint, and Karen P. Scott. 2015. "16S rRNA Gene-Based Profiling of the Human Infant Gut Microbiota is Strongly Influenced by Sample Processing and PCR Primer Choice." *Microbiome* 3 (1): 26.

Walker, Allan. 2010. "Breast Milk as the Gold Standard for Protective Nutrients." *The Journal of Pediatrics* 156 (2): S3-S7.

Walker, A. W., J. Ince, S. H. Duncan, L. M. Webster, G. Holtrop, X. Ze, D. Brown, et al. 2011. "Dominant and Diet-Responsive Groups of Bacteria within the Human Colonic Microbiota." *The ISME Journal* 5 (2): 220-230.

Wall, Steven J. and Dylan R. Edwards. 2002. "Quantitative Reverse Transcription–Polymerase Chain Reaction (RT-PCR): A Comparison of Primer-Dropping, Competitive, and Real-Time RT-PCRs." *Analytical Biochemistry* 300 (2): 269-273.

Walter, J., C. Hertel, G. W. Tannock, C. M. Lis, K. Munro, and W. P. Hammes. 2001. "Detection of Lactobacillus, Pediococcus, Leuconostoc, and Weissella Species in Human Feces by using Group-Specific PCR Primers and Denaturing Gradient Gel Electrophoresis." *Applied and Environmental Microbiology* 67 (6): 2578-2585.

Wang, X., SP Heazlewood, DO Krause, and THJ Florin. 2003. "Molecular Characterization of the Microbial Species that Colonize Human Ileal and Colonic Mucosa by using 16S rDNA Sequence Analysis." *Journal of Applied Microbiology* 95 (3): 508-520.

Wang, Yunwei, Jeanette D. Hoenig, Kathryn J. Malin, Sanaa Qamar, Elaine O. Petrof, Jun Sun, Dionysios A. Antonopoulos, Eugene B. Chang, and Erika C. Claud. 2009. "16S rRNA Gene-Based Analysis of Fecal Microbiota from Preterm Infants with and without Necrotizing Enterocolitis." *The ISME Journal* 3 (8): 944-954.

Weissensteiner, Thomas, Tania Nolan, Stephen A. Bustin, Hugh G. Griffin, and Annette Griffin. 2003. *PCR Technology: Current Innovations* CRC press.

Wesolowska-Andersen, Agata, Martin Iain Bahl, Vera Carvalho, Karsten Kristiansen, Thomas Sicheritz-Pontén, Ramneek Gupta, and Tine Rask Licht. 2014. "Choice of Bacterial DNA Extraction Method from Fecal Material Influences Community Structure as Evaluated by Metagenomic Analysis." *Microbiome* 2 (1): 19.

Yagi, Norio, Kazuhiro Satonaka, Mitsuzo Horio, Hiroyoshi Shimogaki, Yoshio Tokuda, and Sakan Maeda. 1996a. "The Role of DNase and EDTA on DNA Degradation in Formaldehyde Fixed Tissues." *Biotechnic & Histochemistry* 71 (3): 123-129.

Yatsunenko, T., F. E. Rey, M. J. Manary, I. Trehan, M. G. Dominguez-Bello, M. Contreras, M. Magris, et al. 2012. "Human Gut Microbiome Viewed Across Age and Geography." *Nature* 486 (7402): 222-227.

Zhernakova, A., A. Kurilshikov, M. J. Bonder, E. F. Tigchelaar, M. Schirmer, T. Vatanen, Z. Mujagic, et al. 2016. "Population-Based Metagenomics Analysis Reveals Markers for Gut Microbiome Composition and Diversity." *Science (New York, N.Y.)* 352 (6285): 565-569.

Zoetendal, E. G., A. von Wright, T. Vilpponen-Salmela, K. Ben-Amor, A. D. Akkermans, and W. M. de Vos. 2002. "Mucosa-Associated Bacteria in the Human Gastrointestinal Tract are Uniformly Distributed Along the Colon and Differ from the Community Recovered from Feces." *Applied and Environmental Microbiology* 68 (7): 3401-3407.

Zumbo, Paul. "Ethanol Precipitation." Weill Cornell medical college laboratory of Christopher E. Mason, ph.d. Department of physiology & biophysics, accessed September 18, 2017, http://physiology.med.cornell.edu/faculty/mason/lab/zumbo/files/ETHANOL_PRECIPITATION.pdf.

Appendix A- Instruction for Infant's fecal sample collection using wooden coffee stirrer stick

Instruction for infant's fecal sample collection using wooden coffee stirrer stick

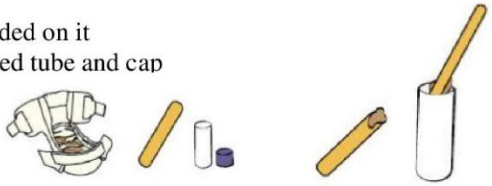
Note: This method is best fit for high throughput fecal sample processing system

Important
- Please read thoroughly before starting to collect feces
- Sarstedt storage container must be stored in freezer (~ -20 °C) with Styrofoam box at least 3 hours before starting the collection
- sample can be collected directly from diaper or pot (avoid urine)

To collect feces you will need:
- Cold storage box (do not take it out from freezer now)
- 2 ml sample tubes with outer grips
- Wooden coffee stirrer sticks
- Labelling paper and a waterproof permanent pen
- Paper plates
- Cotton swab
- Scissor

Start collecting feces after here

Step 1: Preparation
- Put diaper or pot on a flat surface
- Take one paper plate and place all the things needed on it
- Open the cap of 2 ml sample tube and keep opened tube and cap on paper plate



Step 2: Collection (type of feces?)

A: If feces is fluffy and liquid
- Keep the opened tube near to feces and use stick to collect feces at least 4 times into the tube
- Put the wooden stick on paper plate carefully and seal the tube with the cap (in case there is some fecal material on tube, please wipe it before sealing)

B: If feces is firm and hard
- Use wooden stick to scratch out pea sized fecal material and transfer it into the tube
- Put the wooden stick on paper plate carefully and seal the tube with the cap (in case there is some fecal material on tube, please wipe it before sealing)

C: If feces is soaked up by diaper
- Take cotton swab and roll the swab on the area where it moist and more fecal material and carefully slide the swab into the tube and try to break the swab stick (if swab stick does not break than use scissor to cut the tp of swab stick)
- Put the swab on paper plate carefully and seal the tube with the cap (in case there is some fecal material on tube, please wipe it before sealing)

Step 3: Storage
- Take out the storage container from the freezer, keep styrofoam box aside, open the cap in the storage box, insert the tube with collected fecal material into the box, close the cap, keep the box into the styrofoam box and put the box in freezer again
- You can store the sample in the box unless you take it to collection center or someone picks up from your house

Step 4: Disposal
- Dispose used wooden stick, swab stick, paper plate and diaper correctly (remember wooden sticks and swab sticks have sharp edges)

Appendix B- Extraction of Nucleic Acid from 96 fecal samples using Repeated Bead Beating (RBB) with KingFisher Flex 96 (High throughput Protocol)

Using Ambion Bead Tubes and following Magmax™ Protocol

Note: Use Ice Box to keep the samples in between the various steps, do not keep them in room temperature for longer time

*** Reagents available in Kit (**Ambion, MagMax™ Total Nucleic Acid Isolation Kit, AM1840)

Machine: KingFisher Flex, ThermoScientific™

Read all the instruction thoroughly before starting

A. Preparation of Lysis (RBB) buffer

Stock	Dilution (fold)	Final conc.	Amount (ml)
1M Tris-HCl	20	0.05M	5
0,5M EDTA	10	0.05M	10
5M NaCl	10	0.5M	10
20% SDS	5	4%	20
Water (MQ)	-	-	55
Total			100

Prepare stock solutions beforehand

RBB required: 635 µL per sample (**61 ml** for 96 samples)

Preparation of RBB: Mix well until clear solution is observed (May require hot water (max 70^0C) bath to warm the solution) and Store RBB buffer in room temperature.

Usage: If solution precipitates warm the solution in water bath (max 70^0C) to get clear solution

B. Preparation of Nucleic Acid (NA) Bead mix* for final extraction

Each DNA isolation reaction requires **20 µL** of Bead Mix per sample. Although the mixture is stable at 4°C for up to 2 weeks, we recommend preparing it on the day it will be used.

Add 360 µl Carrier NA* (Comes with kit) or Nuclease Free water to the 28 ml Lysis/Binding Enhancer* before mixing with NA Binding Beads*. Lysis/Binding Enhancer is now stable for month and can be stored at room temperature.

Component	Per reaction (µL)	96 rxns (µL)+ 1 extra rxn
NA Binding Beads* (Vortex before using)	10	970
Lysis / Binding Enhancer*	10	970
Total	20	1940

The bead mix can be stored at 4°C until used (Vortex well before using)

Protocol continues to next page

C. Preparation of 1X PBS pH 7.4 (Note: Readymade 1X PBS can also be used from Kit)

Stock	Dilution (fold)	Final conc.	Amount (ml)
10X PBS	10	1X	10
Water (MQ)	-	-	90
Total			100

Mix well and store at room temperature

D. Isopropanol (No need to prepare Just ensure the required amount before starting the extraction)

Isopropanol is required for the final NA (Nucleic Acid) extraction process and **300 µL is required per sample.**

E. Preparation of Wash Buffers for DNA purification

(Wash Buffers should be ready before starting extraction process)

To prepare wash buffers: Go to section H

Wash Buffers constituents			
Wash Buffer 1			
HWI_E			
	Stock	**Final**	**Volume (ml)**
Na-acetate pH 5.2	3M	0.3M	6.5
EDTA	0.5M	10mM	1.3
Ethanol	96 %	70 %	57.2
Total			65
Wash Buffer 2			
HWII			
	Stock	**Final**	**Volume (ml)**
Tris-HCl pH 8	1M	10mM	1
Ethanol	96 %	70 %	99
Total			100

Note: If Wash buffers precipitates prepare again

Protocol continues to next page

F. Disruption of Solid Samples (Feces sample) and collecting supernatants (Follow the steps from 1 to 4)

1. Weigh ~0.25 g sample (Feces) and mix with PBS

a. Weigh ~0.25 g (±0.05 g) sample to the 2ml Eppendorf Tube

b. Pipette **1 mL** of 1X PBS (pH 7.4) into the sample tube for Adult feces samples and **0.5 ml** 1X PBS (pH 7.4) for baby feces samples

c. Mix by vortexing (vertically and horizontally) at moderate to high speed until the sample is fully suspended (Homogenized)

Note: The Vortemp 56 (Incubator) can only fit 56 sample tubes so it's wise to process 48 (half plate) samples for first round of bead beating followed by preparation of another set of 48 samples.

3. Centrifuge at 100 x g for 1 min

a. Centrifuge fecal samples (Homogenized with PBS) at very low speed, 100 x g (971 rpm), for 1 min

(Note: step 2 and 3 can be done at the same time)

b. Transfer **175 µL** of the supernatant into each Bead Tube **(Step 1)** containing RBB Buffer Solution (Total volume **410 µL**) Continue to section F 4.0 and **step IV**

2. Add 235 µL Lysis buffer (RBB) to Beaded Tubes

Dispense 235 µL RBB Buffer into Beaded Tubes (Ambion) with fecal supernatant (From Step 2) for each sample.

Perform the Repeated Bead Beating

Protocol continues to next page

4. Perform Repeated Bead Beating

Note: Be prepared with well labelled KingFisher deep well 96 plate for collection of supernatants (S1 and S2) and meanwhile set the water bath or Vortemp 56 (labnet) at 95^0C for the incubation before starting the bead beating.

Follow the following procedure for repeated bead beating (RBB)

Steps	Procedure
I	Use FastPrep®-96 (MP) at speed 800 rpm for 60 seconds (After 60 second run let samples and machine cool down for 5 min, Samples can remain at Room Temperature) and repeat the fast prep for **2 more times**
II	Incubate the tubes at 95°C for 15 min with gentle shaking (Vortemp 56)
III	Centrifuge at RT for 5 min at full speed (to pellet the solid particles) and Transfer **~300 ul** supernatant **(S1)** into deep well 96 plate **(Remember: avoid pipetting beads into the supernatants)** and seal the plate unless required.
IV	Add 100µL of fresh lysis (RBB) buffer to the same bead beating tube with pelleted feces and repeat the whole process from **Step I to III** and Transfer **all (Final)** the supernatant **(S2) (Use same deep well 96 plate to collect both supernatants)**
V	Collect both Supernatants **(S1 and S2)** in same plate and seal the plate unless required for purification. Plate can remain in room temperature at least for 30 minutes while preparing the plates for purification. **Important note**: If purification cannot be started immediately, plate with supernatants can be sealed and stored in +4°C until plate is required for purification process and due to the precipitative nature of RBB buffer when placed long in room temperature or in cold temperature it is highly recommended to start the purification process as soon as both supernatants are collected. If supernatants start to form pellets transfer the deep well plate into the warm water bath to resuspend the supernatants completely.

Follow the Automated Nucleic acid Extraction protocol using KingFisher FLEX, ThermoScientific (Section G)

Protocol continues to next page

G. Automated Nucleic acid Extraction protocol using KingFisher FLEX, ThermoScientific

Manual: https://www.thermofisher.com/us/en/home/life-science/dna-rna-purification-analysis/automated-nucleic-acid-purification/kingfisher-nucleic-acid-protein-purification-systems/kingfisher-flex-system.html
***Kit:** Ambion (life technologies) MagmaxTM Total Nucleic Acid Isolation Kit

Follow the following steps to extract the Nucleic acid from Bead beated samples

Preparing Plates For extracting and purifying 96 samples

Note: Prepare following plates 10-20 minutes before transferring them to the machine

Pipette the reagents into the 96 well plates as indicated below using MultiDrop Combi, ThermoScientific

Plate type	Plate code	Samples and reagents	VOLUME (µl)
Deep well	**WashI_1**	**1st wash I**	
		HWI_E	300
Deep well	**WashI_2**	**2nd Wash I**	
		HWI_E	300
Deep well	**WashII_1**	**1st Wash II**	
		HWII	450
Deep well	**WashII_2**	**2nd Wash II**	
		HWII	450
Standard	**Elution (E)**	Elution Buffer*	50
Standard	**TIP**	TIP (Just put on the standard plate)	
Deep well	**Sample**	Bead Mix* (pipette manually)	20
		Isopropanol	300
		RBB	300

***Reagents used from Kit**

Note: Label the elution plate carefully since the extracted Nucleic acid can be directly stored into -20^{0}C

Prepare the Sample plate for the machine:
Mix and transfer **200 µL** of the supernatants (S1 and S2) collected in deep well 96 plate after bead beating in **step 4** into the sample plate prepared above **(plate code: Sample).**

All required plates are prepared at this step

Turn the machine (KingFisher FLEX™) on, take off the cover and carefully change the magnet from maintenance screen.

Select the protocol for the extraction from home screen, in our case protocol is **MagMAX Pathogen High Vol Duo** and start the program. Place the plates according to the instruction displayed on the screen and run the program. **Note**: It's a **31 min** protocol after program has started.

When program ends seal the Elution Plate carefully and store at +4^{0}C if concentration is measured in same day or next day. **Note: If concentration will not be measured next day, store the plate in -20^{0}C**. Measure the concentration and store the samples in same plate or transfer the samples into other 96 well plate or strips and store in -20^{0}C.

Protocol continues to next page

H. Wash buffers Preparation

Prepare wash buffer 1 with Unautoclaved reagents			
HWI_E			**For 100 ml**
	Stock	**Final**	**Volume (ml)**
Na-acetate pH 5.2	3M	0.3M	10
EDTA	0.5M	10mM	2
Carefully maintain pH 7.5 to 8.0 using 10M NaOH (~3ml required, pH might jump after adding >2ml NaOH)			
Send to autoclave and carefully adjust final volume (100 ml) by adding 70% EtOH)			

Wash Buffer II (Prepare with autoclaved Tris-HCl)			
HWII			**For 100 ml**
	Stock	**Final**	**Volume (ml)**
Tris-HCl (From Stock prepared for LYSIS buffer)-	1M	10mM	1
Ethanol	96 %	70 %	99
Total			**100**

Protocol ends here